

Political deepfakes and the new laws in NSW

Lenny Roth BCom, LLB

Senior Research Officer, Parliamentary Research Service

Tom Gotsis BA, LLB, Dip Ed, Grad Dip Soc Sci

Research Officer, Parliamentary Research Service

April 2026



Key points

- Political deepfakes are AI-generated digital content that depict a politician doing or saying something they did not do or say. They can also depict someone falsely expressing a view about a politician or political issue.
- Political deepfakes can be deceptive, sexually explicit or satirical. They can also draw positive and negative associations through imagery, analogies and associations.
- Since 2017 one database has captured 2,201 global incidents of political deepfakes. Another database recorded 453 global incidents of political deepfakes in 2025.
- Political deepfakes may have adverse effects on elections, politicians and political communication.
- In March 2026 the NSW Government introduced electoral law reforms including a ban on publishing deepfakes during an election period.
- There was debate in parliament as to whether the new offences are consistent with the implied freedom of political communication in the Australian Constitution.
- Other jurisdictions have also enacted political deepfake laws. These include South Australia, many states in the United States, and one Canadian province.
- Other measures for combating political deepfakes include the use of watermarks, deepfake detection technology, removal of deepfakes by social media platforms, and public education and awareness.
- 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 and democracy in NSW.

Contents

Key points.....	1
1. Introduction	3
2. What are political deepfakes?	4
3. Prevalence.....	6
3.1 Actual prevalence.....	6
3.2 Perceived prevalence	8
4. Potential adverse impacts.....	9
4.1 Elections.....	9
4.2 Effects on politicians	11
4.3 Broader effects on political communication.....	14
5. Law reform in New South Wales.....	16
5.1 New electoral law offences	16
5.2 Constitutional issues	17
5.3 Other laws in NSW.....	18
6. Law reform in other jurisdictions.....	19
6.1 Australia	19
6.2 International.....	20
7. Other responses to political deepfakes	22
7.1 Deepfake counter measures	22
7.2 Actions taken by digital platforms	24
7.3 Actions taken by electoral authorities	25
8. Conclusion.....	26

1. Introduction

There are international concerns about the potential impacts of artificial intelligence (AI) on democracy.¹ Among these concerns is the proliferation of deepfakes that are designed to mislead voters around elections.² A deepfake is AI-generated digital content that depicts a person saying or doing things they did not actually say or do. Deepfakes have 'been deployed to influence elections in India, Indonesia, Mexico, Pakistan, Slovakia, Ukraine, the USA and Taiwan.'³ 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 aims to provide background to, and a summary of the NSW legislative reforms. It begins by outlining the problem of political deepfakes, including their prevalence and potential adverse impacts on elections, politicians and political communication. The paper then outlines the NSW reforms, including constitutional issues. Next, the paper reviews legislative developments in other Australian jurisdictions and overseas. Finally, other measures that can be used to combat political deepfakes are discussed. Broader topics that are related to but not the focus of this paper include the regulation of truth in political advertising⁴ and online misinformation.⁵

¹ C Summerfield et al, [The impact of advanced AI systems on democracy](#), *Nature Human Behaviour*, 2025, 9: 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 et al, [The impact of advanced AI systems on democracy](#), *Nature Human Behaviour*, 2025, 9: p 2423. 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.

³ C Summerfield et al, [The impact of advanced AI systems on democracy](#), *Nature Human Behaviour*, 2025, 9: p 2423-2424.

⁴ This issue was discussed in Joint Standing Committee on Electoral Matters, [Administration of the 2023 NSW state election and other matters](#), Parliament of NSW, October 2024, Ch 2.

⁵ See N Fraser, [What's next for misinformation regulation?](#), Commonwealth Parliamentary Library, Flagpost, 2 July 2025.

2. What are political deepfakes?

A 2024 systematic literature review found that there is as yet no universally accepted definition of the term 'deepfake'.⁶ 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.⁷ For the purposes of this paper, political deepfakes are defined as deepfakes that depict: (a) a politician, or (b) a person expressing a view about a politician or political issue.⁸

Political deepfakes most typically take the form of videos, images and audio, or a combination of those forms. However, as illustrated by the British politician who 'turned himself into an AI chatbot',⁹ as well as the Czech village with a 'digital mayor',¹⁰ it is possible to create conversational chatbots with the online persona of a politician. These British and Czech examples were done at the instigation of the politicians themselves; however, deepfake chatbots could be used to create a false online persona of a politician that generates fake responses to questions.¹¹ Table 1 provides a technical classification of political deepfakes.

Table 1: Technical classification of political deepfakes

Type	Description
Face-swapped images and videos	A politician's face is overlaid onto another person's body
Lip-syncing and audio overlays	A politician's mouth movements are replaced to match synthetic or manipulated audio
Voice-only cloning	A politician's voice is replicated and can be used to say anything
Full-body re-enactment	An actor's posture, movement and gestures can be used to make a politician seem to be acting in ways that they never did
Conversational chat-bots	Can be used to create fake online personas of politicians that give misleading spoken or written responses to questions

Based on: Sentinel One, [Deepfakes: Definition, Types and Key Examples](#), 16 July 2025.

In addition to classifying political deepfakes based on their technical characteristics, political deepfakes can also be classified in terms of their content. One content-based approach classifies

⁶ 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*, 2024.

⁷ See also: T Gotsis, [Sexually explicit deepfakes and the criminal law](#), NSW Parliamentary Research Service, 2025, p 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.

⁸ For instance, a deepfake has depicted fake fans of Taylor Swift supporting Donald Trump, while an anti-immigration deepfake has depicted fake Haitian immigrants rounding up and eating pets in Ohio: R Youngblood, [I knew you were trouble: Deepfakes, misinformation and the threat to elections](#), *Louisiana Law Review*, 2025, 86, 320, 331-334. Bollywood actors have also been falsely depicted as criticising Prime Minister Narendra Modi and asking people to vote for the opposition Congress party in an upcoming election: [Deepfakes of Aamir Khan and Ranveer Singh raise worries over AI misuse in Lok Sabha election 2024](#), *Hindustan Times*, 22 April 2024.

⁹ R Min, [A British politician turned himself into an AI chatbot. Meet the UK's first 'virtual MP'](#), *Euro News*, 6 August 2025.

¹⁰ D Lazarova and G Hykl, [Czech village wows country with first digital mayor](#), *Radio Prague International*, 10 April 2026.

¹¹ The Australian Electoral Commission is concerned about chatbots providing false information about election processes: Australian Electoral Commission, [AI & Elections](#), AI communication channels, Chatbots, 25 September 2025.

political deepfakes on the basis of how realistic they are and whether their sentiment is positive or negative.¹² This classification results in the 4 types of political deepfakes shown in Table 2.

Table 2: Content-based classification of political deepfakes

Type	Features
Potentially deceptive and negative	<ul style="list-style-type: none"> Pose a direct threat to a politician’s wellbeing and reputation For example, they may show a politician being sexually explicit, violent, drunk, corrupt, inept, rude, or endorsing hate speech or violence They have the potential to sway voters through deliberate deception because they are extremely realistic
Potentially deceptive and positive	<ul style="list-style-type: none"> Aim to glorify a politician through fabricated positive moments For example, they may show a politician deliver a powerful speech they never gave or do something heroic, virtuous or charitable, or battling adversity They have the potential to sway voters through deliberate deception because they are extremely realistic
Not intending to deceive and negative	<ul style="list-style-type: none"> Depict a politician negatively using obviously fictional scenarios (or are clearly marked as being AI-generated) For example, a politician may be depicted as a fictional character (such as the grim reaper) or as appearing at a historical event (such as a Nazi rally) Often described as a form of political satire Potential to sway voters does not depend on extreme realism or deliberate deception, but on satire, analogies and negative associations
Not intending to deceive and positive	<ul style="list-style-type: none"> Depict a politician positively in an obviously fictional scenario (or are clearly marked as being AI-generated) For example, a politician may be shown as a superhero or revered historical figure Seek to develop a cult of personality 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

Based on: M Wack et al, [Scrutinizing the many faces of political deepfakes](#), *Tech Policy Press*, 18 November 2025.

As this content-based typology suggests, political deepfakes can include attempts to intimidate or discredit a politician by falsely depicting them in sexually explicit material, making them appear corrupt or inept, or attributing comments to them that are offensive or opposed to their values. They can be satirical in nature. Political deepfakes can also be used to enhance a politician’s own standing through glorification and deception.

¹² M Wack et al, [Scrutinizing the many faces of political deepfakes](#), *Tech Policy Press*, 18 November 2025 and C Walker et al, [Beyond deception: A functional typology of political deepfakes](#), SSRN, n.d., p 20, accessed 17 February 2026.

3. Prevalence

3.1 Actual prevalence

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.¹³ However, as discussed below (see 3.1.1 and 3.1.2), both research institutions and private companies are attempting to identify the number and type of deepfakes that are online and track future trends.

3.1.1 The Political Deepfakes Incidents Database

In 2023 a university and private sector collaboration started collecting data for a new database, the [Political Deepfakes Incidents Database](#).¹⁴ The database records incidents of deepfakes that are assessed as having social or political significance. It also tracks false allegations of deepfakes (that is, where a genuine photo of an incident was falsely claimed to be a deepfake by the perpetrator of the incident). The database includes data from 2017 onwards. When it launched on 26 January 2024, the database included details of 114 political deepfakes.¹⁵ As at 20 April 2026, the database included records of 2,201 political deepfakes.¹⁶

One Australian political deepfake was located in the database. It was an image of Georgie Purcell, a member of the Victorian Parliament, which was published by a media outlet after it sexualised her image by making her clothing more revealing and enlarging her breasts.¹⁷

The establishment of the database is an important development. Figures from the database should, however, be treated cautiously, as they are likely to underrepresent the actual number of political deepfakes that are online. This is because only deepfakes with 'social or political significance' are included in the database, which introduces a subjective element to the process, and because the data capture process is unlikely to be perfect.¹⁸ Demonstrating the point, there are other Australian political deepfakes that have been reported in the media which do not appear in the database, including:

¹³ 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*, 2024.

¹⁴ A Azzo, [Tracking political deepfakes: New database aims to inform, inspire policy solutions](#), Centre for Advancing Safety of Machine Intelligence, 26 January 2024.

¹⁵ A Azzo, [Tracking political deepfakes: New database aims to inform, inspire policy solutions](#), Centre for Advancing Safety of Machine Intelligence, 26 January 2024.

¹⁶ This excludes 135 instances where it was unclear whether material was a deepfake or a cheapfake, and 13 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.

¹⁷ 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.

¹⁸ A Azzo, [Tracking political deepfakes: New database aims to inform, inspire policy solutions](#), Centre for Advancing Safety of Machine Intelligence, 26 January 2024.

- A deepfake of former Queensland Premier Anastacia Palaszczuk that was circulated during the 2020 state election period in which she described Queensland as having ‘massive debt’ and ‘huge unemployment’¹⁹
- A deepfake of then federal opposition leader Peter Dutton dancing in support of nuclear power plants²⁰
- Deepfakes of Prime Minister Albanese and then opposition leader Peter Dutton that were made by Senator David Pocock to illustrate the ease with which deepfakes can be created.²¹

A 2025 academic paper used the Political Deepfakes Incidents Database to gather a sample of 70 political deepfakes from the 2024 US presidential election.²² When the classification scheme shown in Table 2 was applied to the sample, it was found that:

- 43.5% of the deepfakes were non-deceptive and positive
- 34.8% of the deepfakes were non-deceptive and negative
- 11.6% of the deepfakes were potentially deceptive and negative
- 10.1% of the deepfakes were potentially deceptive and positive.²³

The 2025 paper noted that 78.3% of the political deepfakes sought to sway political sentiment not through deception but through such factors as emotional resonance, humour, cultural reference and rhetorical framing.²⁴ Deception was the basis of the remaining 21.7% of deepfakes.

3.1.2 Resemble AI Threat Landscape Briefing: January 2025 to January 2026

Resemble AI²⁵ provided the Parliamentary Research Service with a Threat Landscape Briefing based on data from its own Deepfake Incident Database. Resemble AI identified 1,722 global deepfake incidents between 1 January 2025 to 31 December 2025 (Figure 1). Of those 1,722 incidents, the largest victim category was private citizens (616 or 36%),²⁶ followed by public figures (456 or 26%). These categories were closely followed by political deepfakes (453 or 26%), and then enterprises (197 or 11%). Resemble AI noted that over the 12 months the number of political deepfakes peaked at 65 in February 2025.

¹⁹ S Richards, [As deepfakes advance with technology, there are concerns they could become a ‘threat to democracy’](#), ABC News, 8 October 2024. It was claimed that these deepfakes were satire because they were clearly marked as being deepfakes. See also: Queensland Government, [Disinformation and elections in the age of artificial intelligence](#), 29 February 2024.

²⁰ S Richards, [As deepfakes advance with technology, there are concerns they could become a ‘threat to democracy’](#), ABC News, 8 October 2024. It was claimed that these deepfakes were satire because they were clearly marked as being deepfakes.

²¹ J Evans, [Senator David Pocock creates AI deepfakes of Anthony Albanese and Peter Dutton to call for ban ahead of election](#), ABC News, 7 September 2024.

²² C Walker et al, [A Beyond deception: A functional typology of political deepfakes](#), SSRN, 3 November 2025, p 20-21, and M Wack et al, [Scrutinizing the many faces of political deepfakes](#), Tech Policy Press, 18 November 2025.

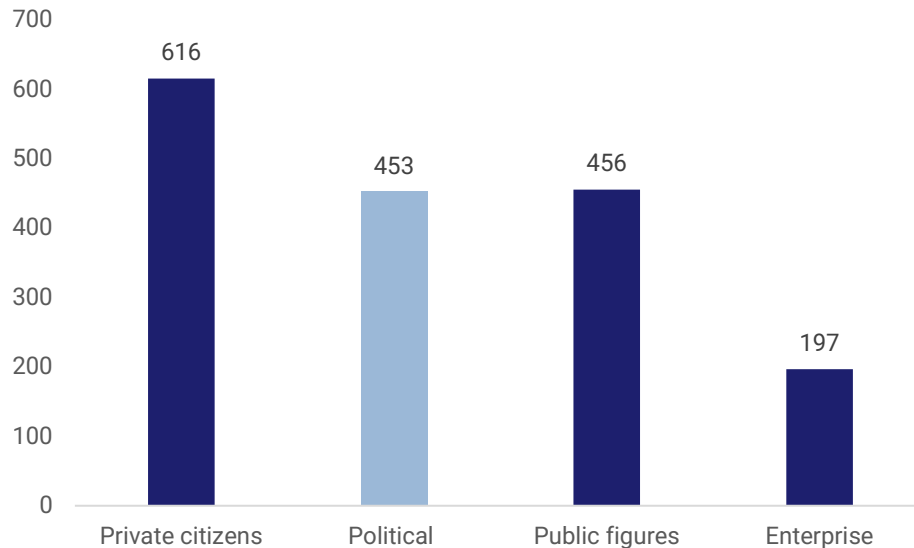
²³ C Walker et al, [A Beyond deception: A functional typology of political deepfakes](#), SSRN, 3 November 2025, p 20-21, and M Wack et al, [Scrutinizing the many faces of political deepfakes](#), Tech Policy Press, 18 November 2025.

²⁴ M Wack et al, [Scrutinizing the many faces of political deepfakes](#), Tech Policy Press, 18 November 2025.

²⁵ [Resemble AI](#) specialises in generative AI voice products, content watermarking and deepfake detection.

²⁶ This category can include financial scams that target individuals, sexually explicit deepfakes and deepfake child abuse material.

Figure 1: Global deepfake incidents by victim category, 2025



Source: Resemble AI

3.2 Perceived prevalence

In March 2025, in the lead up to Australia’s 2025 federal election, software developer Adobe surveyed 1,010 Australians about the impact of deepfakes on elections.²⁷ The survey found that 77% of respondents believed there had been an increase in the number of political deepfakes in the past 3 months, with 25% of respondents believing there had been a significant increase. Most respondents were concerned about the impact of deepfakes on political discourse and elections:

- 86% of respondents believed that AI has made it more difficult to tell whether digital content is real
- 69% of respondents were concerned about the impact of deepfakes on the upcoming federal election
- 78% of respondents said that they supported the introduction of stricter regulations requiring clear labelling of AI-generated political content, in order to protect voters.²⁸

²⁷ Adobe, [Adobe research shows political deepfakes are rising ahead of the Australian federal election](#), 9 April 2025, and Adobe, [Authenticity in the Age of AI: Australia](#), April 2025.

²⁸ Adobe, [Authenticity in the Age of AI: Australia](#), April 2025.

4. Potential adverse impacts

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

4.1 Elections

Some political deepfakes seek to influence people's voting preferences through the use of deception.²⁹ Such deepfakes have the potential to create uncertainty and doubt about all political information that is presented to voters before an election.³⁰ There can also be a national security dimension to this issue, where political deepfakes have been created by a foreign state to interfere in the elections of other states.³¹ 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.³²

The threat to elections and democracy can be seen in the events surrounding the 2025 Irish presidential election. The 2025 Irish presidential election was the subject of 172 instances of political misinformation and disinformation attacks published on X, TikTok, Facebook and YouTube in the 6 weeks leading up to polling day.³³ 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.³⁴ Ms Connolly reported the image to the relevant social media platforms and lodged a formal complaint with the Electoral Commission, after which the deepfakes were removed from some platforms.³⁵ 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.³⁶ Ms Connolly won the presidential election.

²⁹ M Wack et al, [Scrutinizing the many faces of political deepfakes](#), *Tech Policy Press*, 18 November 2025. 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.

³⁰ 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*, December 2025, Vol 27(12), p 6,819-6838.

³¹ 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.

³² A Ray, [Disinformation, deepfakes and democracies: The need for legislative reform](#), *UNSW Law Journal*, 2021, Vol 44(2), p 987. See also: M Platow et al, [Information resilience: How misinformation undermines social cohesion, trust and democracy](#), 2026, report to the National Science and Technology Council, Australian Government Office of the Chief Scientist, Ch 1.4.

³³ Institute for Strategic Dialogue, [Irish Presidential Election 2025: Renewed attacks on election integrity and repeated platform failures](#), 31 October 2025.

³⁴ J Liggett, ['Disgraceful' deep-fake AI video condemned by presidential candidate](#), *BBC*, 23 October 2025.

³⁵ T McNally, ['Malicious deepfake': Condemnation of AI video falsely showing Catherine Connolly ending campaign](#), *Irish Examiner*, 22 October 2025.

³⁶ Connolly for President, [Connolly condemns 'malicious deep-fake' video](#), media release, 22 October 2025, accessed 24 March 2026.

4.1.1 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’³⁷ or election results.³⁸

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.’³⁹ 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.’⁴⁰ It has also been suggested that the lack of clear evidence is ‘no cause for complacency’.⁴¹

While the lack of clear evidence about the impact of political deepfakes may initially appear counterintuitive, on reflection it is not surprising. Elections are complex and dynamic events, with many variables that can complicate efforts to measure the potential impact of political deepfakes on election results. For instance, investigating whether and to what extent political deepfakes influenced voters would require polling.⁴² However, people are typically exposed to vast amounts of information during an election, and many conscious, subconscious and emotional factors can influence how people vote. People may not be able to identify what, if any, information influenced their vote at all, let alone whether they could have been influenced by a deepfake.

Another variable which can complicate efforts to measure the potential impact of political deepfakes on election results is the extent to which any deepfake counter measures were used (see 7.1.1).

4.1.2 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.⁴³

³⁷ O Carpenter-Zehe, [Irish election deepfake AI video shines light on lack of EU-wide rules](#), *euobserver*, 27 October 2025, citing Eva Lejla Podgorsek, Senior Policy Manager at the Non-Government Organisation [Algorithm Watch](#).

³⁸ 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*, 26 April 2024, Vol 23, 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’.

³⁹ Australian Electoral Commission, [AI & Elections](#), 25 September 2025.

⁴⁰ A Ray, [Disinformation, deepfakes and democracies: The need for legislative reform](#), *UNSW Law Journal*, 2021, Vol 44(2), p 988.

⁴¹ 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.

⁴² Polling is already a complex exercise that has been prone to error: see, for instance: T Shepherd, [The trend is in, but Australian voters’ views are soft and fragmented – how should we read the polls?](#) *The Guardian*, 17 April 2025.

⁴³ 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 28.

A 2024 study examined a sample of deepfakes that were reported in 11 countries which had an election in 2023 or had ongoing election campaigns (USA, Turkiye, Argentina, Poland, United Kingdom, France, India, Bulgaria, Taiwan, Indonesia and Slovakia).⁴⁴ The study found that none of the deepfakes which they examined had a decisive impact on the course of elections.⁴⁵ It added that this finding 'does not mean no effect at all'. In particular:

... there are some signs of already noticeable phenomena of an undermined trust in information, politics and the media, which strengthens the sense of uncertainty among the society and opens up new possibilities for manipulation.⁴⁶

The Australian Senate's Select Committee on Adopting Artificial Intelligence found that, while AI has the potential to improve some electoral processes, the available evidence presented to it 'reflects the understanding, both in Australia and globally, that the current state of AI technology brings with it significant risks in relation to the conduct of electoral processes'.⁴⁷ The committee noted:

The risk of AI generated political disinformation is not only that it can be used to influence the outcome of political debates or contests, but also that the uncertainty it creates can lead to an erosion of public trust in and engagement with politics more generally.⁴⁸

The committee noted that the risk posed by deepfakes to elections may vary according to the date on which a deepfake is deployed. Deepfakes deployed in the earlier stages of an election period could seek to undermine the reputation of candidates or shape voter attitudes. Deepfakes deployed close to or on an election day could be better placed to confuse voters about the election process. Finally, deepfakes which are intended to erode confidence in the integrity of an election outcome (by, for instance, allegations of electoral fraud) could be deployed after the date of an election.⁴⁹

4.2 Effects on politicians

4.2.1 Reputational harm

Concerns have been raised for some time about the scope for political deepfakes to cause reputational harm to politicians.⁵⁰ It has even been suggested 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.⁵¹

⁴⁴ M Labuz and C Nehring, [On the way to deepfake democracy? Deepfakes in election campaigns in 2023?](#), *European Political Science*, 26 April 2024, Vol 23, 454–473.

⁴⁵ M Labuz and C Nehring, [On the way to deepfake democracy? Deepfakes in election campaigns in 2023?](#), *European Political Science*, 26 April 2024, Vol 23, 454–473.

⁴⁶ M Labuz and C Nehring, [On the way to deepfake democracy? Deepfakes in election campaigns in 2023?](#), *European Political Science*, 26 April 2024, Vol 23, 454–473 at 468. The authors noted, however, that this 'does not mean no effect at all'.

⁴⁷ Australian Parliament, the Senate, [Select Committee on Adopting Artificial Intelligence \(AI\): Interim Report](#), October 2024, p 29.

⁴⁸ Australian Parliament, the Senate, [Select Committee on Adopting Artificial Intelligence \(AI\): Interim Report](#), October 2024, p 9.

⁴⁹ Australian Parliament, the Senate, [Select Committee on Adopting Artificial Intelligence \(AI\): Interim Report](#), October 2024, p 14.

⁵⁰ See, for instance, B Chesney and Danielle Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 2019, Vol 107, 1753-1820, p 1774.

⁵¹ 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 8 and 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.

The empirical evidence on this topic is only starting to emerge and is experimental in nature, rather than based on real-world outcomes, such as polling or election results. A 2024 study showed that deepfakes can cause reputational harm to politicians and do not have to be particularly credible to do so.⁵² A 2025 study showed that political deepfake videos which suggested a sex, corruption or prejudice scandal resulted in 'substantial reputational damage' for innocent politicians.⁵³ An important caveat to that finding was that journalistic fact checking could substantially reduce or even eliminate the reputational damage.

4.2.2 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.⁵⁴ 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.⁵⁵ Moreover, deepfakes can also create 'a form of reality–identity dissonance, where the lifelike but fabricated portrayal destabilises the individual's sense of self and perceived social identity'.⁵⁶

A particularly harmful form of deepfake victimisation involves sexualised and sexually explicit deepfakes being used to intimidate, demean and blackmail politicians, particularly (although not exclusively) female politicians.⁵⁷ The literature indicates that victims of sexually explicit deepfakes suffer from emotional, psychological, personal, social and professional distress.⁵⁸ Several female politicians in the United Kingdom have already been targeted in this way, with one victim noting that 'none of this is about sexual pleasure, it's all about power and control'.⁵⁹ In Malaysia 10 politicians, including at least 4 male politicians, were blackmailed with sexually explicit deepfakes, with the offenders demanding \$US100,000 in cryptocurrency.⁶⁰

⁵² M Hameleers, T van der Meer and T Dobber, [Distorting the truth versus blatant lies: The effects of different degrees of deception in domestic and foreign political deepfakes](#), *Computers in Human Behaviour*, March 2024, Vol 152 at 11. Implausible deepfakes (those which deviated considerably from a politician's known policy position) were shown to cause greater reputational harm than more plausible deepfakes.

⁵³ V Dan, [Deepfakes as a democratic threat: Experimental evidence shows noxious effects that are reducible through journalistic fact checks](#), *The International Journal of Press/Politics*, 11 February 2025.

⁵⁴ M Umar, [Digital trauma: deepfake victimization and AI-generated violence](#), *The Lancet: Psychiatry*, February 2026, Vol 13(2), p 89-90.

⁵⁵ M Umar, [Digital trauma: deepfake victimization and AI-generated violence](#), *The Lancet: Psychiatry*, February 2026, Vol 13(2), p 89-90.

⁵⁶ M Umar, [Digital trauma: deepfake victimization and AI-generated violence](#), *The Lancet: Psychiatry*, February 2026, Vol 13(2), p 89-90.

⁵⁷ 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.

⁵⁸ 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*, 2021. See also: T Gotsis, [Sexually explicit deepfakes and the criminal law](#), NSW Parliamentary Research Service, 2025, p 8-10.

⁵⁹ See, for instance: J Waterson, [British female politicians targeted by fake pornography](#), *The Guardian*, 2 July 2024.

⁶⁰ [Malaysian politicians targeted in AI deepfake blackmail campaign](#), *Folio3*, September 2025. See also: M Mustafa, [Malaysia grapples with AI legal grey zone as deepfake porn blackmail targets lawmakers](#), *The Straits Times*, 23 September 2025.

In Australia, as discussed earlier (see 3.1.1), Victorian MP Georgie Purcell had a sexualised image of her published by the media. Ms Purcell referred to the incident as a form of discrimination and objectification, noting that ‘... this is not something that happens to my male colleagues ...’.⁶¹ She added that the ‘... message this sends to young women and girls across Victoria is that even at the top of your field, your body is always up for grabs.’⁶²

4.2.3 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.

Women face particular risks from political deepfake victimisation. A 2025 article found that 26 members of Congress were depicted in more than 35,000 instances of deepfake content on pornographic websites, and 25 of these members were women.⁶³ The article stated that this:

... equates to nearly one in six women in Congress becoming victims of this insidious technology. Such attacks ... reveal how AI tools are being used to threaten women in leadership, potentially deterring their political participation.⁶⁴

The 2025 article also referred to cases of deepfake victimisation of female politicians that have occurred in the United Kingdom, Italy and Pakistan.⁶⁵

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.⁶⁶

4.2.4 Public opinion of constituents may be harder to discern

Online environments can include fake communication from constituents, which makes it challenging for politicians to determine what constitutes genuine public opinion in their electorates.⁶⁷ Deepfakes and ‘bots’⁶⁸ can be used to generate fake emails, messages, correspondence and articles on mass to create the illusion of public opinion.⁶⁹

⁶¹ 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.

⁶² 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.

⁶³ M Clarice, [Deepfakes, AI, and the new war against women in politics](#), *WePro*, 8 January 2025.

⁶⁴ M Clarice, [Deepfakes, AI, and the new war against women in politics](#), *WePro*, 8 January 2025. See also: L O’Neill, [Fake photos, real harm: AOC and the fight against AI porn](#), *Rolling Stone*, 8 April 2024.

⁶⁵ M Clarice, [Deepfakes, AI, and the new war against women in politics](#), *WePro*, 8 January 2025.

⁶⁶ Inter-Parliamentary Union, [Dangers of deepfakes for parliamentarians](#), 19 February 2024.

⁶⁷ M Adam and C Hocquard, [Artificial intelligence, democracy and elections](#), European Parliamentary Research Service, October 2023, p 3-4.

⁶⁸ J Watts, [What exactly is a social media bot?](#) White Space, 22 August 2025, accessed 13 March 2026 and M Gigashvili, [What are social media bots?](#), Medium, 9 May 2025, accessed 13 March 2025.

⁶⁹ F Menczer, [Swarms of AI bots can sway people’s beliefs – threatening democracy](#), *The Conversation*, 13 February 2026 and M Adam and C Hocquard, [Artificial intelligence, democracy and elections](#), European Parliamentary Research Service, October

4.3 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.⁷⁰ 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.'⁷¹

Australia's political history has frequently shown that building a foundation of shared facts and truths can be challenging.⁷² 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.⁷³ Academics have suggested that political deepfakes have the potential to further exacerbate those challenges, which will increase political polarisation,⁷⁴ confuse people as to what constitutes a fact, and:

... 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.⁷⁵

Political deepfakes could also 'contribute to a general climate of uncertainty and doubt', which would reduce overall trust in political information.⁷⁶ The climate of uncertainty and doubt created by deepfakes further damages political discourse by making it easier for politicians to deny the truth. Academics 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

2023, p 4. Adam and Hocquard note that 'An experiment showed that legislators found AI-generated text they received on six policy areas almost as credible as human-written messages.' See: S Kreps and D Kriner, [How generative AI impacts democratic engagement](#), *Brookings*, 21 March 2023.

⁷⁰ Note that 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.

⁷¹ B Chesney and D Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 2019, Vol 107, 1753-1820 at 1,777.

⁷² 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.

⁷³ B Chesney and D Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 2019, Vol 107, 1753-1820, at 1764–1768 and 1786. See also: M Platow et al, [Information resilience: How misinformation undermines social cohesion, trust and democracy](#), 2026, report to the National Science and Technology Council, Australian Government Office of the Chief Scientist.

⁷⁴ 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*, 2025, Vol 44, p 321-339 (links to abstract only).

⁷⁵ B Chesney and D Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 2019, Vol 107, 1753-1820, p 1778.

⁷⁶ 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*, December 2025, Vol 27(12), p 6,819-6838.

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 actually fake.⁷⁷

It has been argued that the rise in political deepfakes can be seen an extension of the 'fake news' phenomenon, where claims of 'fake news' are used to deny the truth of facts, avoid accountability and bypass rational debate. It has also been claimed that there is a financial dimension to this issue; namely, that political deepfakes are being posted on fake news channels for profit, either through advertising or through social media companies paying content creators based on levels of user engagement.⁷⁸

When the ability to know what is true is diminished there is a risk that the trust which is necessary for democracy to function effectively will be eroded.⁷⁹ In such a climate, people's ability to be included in political debates and decisions that affect them is reduced, as is the legitimacy of political decisions.⁸⁰ Moreover, the 'combination of *truth* decay and *trust* decay ... creates greater space for authoritarianism' because ... 'leaders with authoritarian tendencies benefit when objective truths lose their power.'⁸¹

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'.⁸² Political deepfakes which are satirical, rather than deceptive, arguably make a distinct and significant contribution to political discourse.⁸³ 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'.⁸⁴

⁷⁷ B Chesney and D Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 2019, Vol 107, 1753-1820, 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 17 March 2026.

⁷⁸ M Workman et al, [Foreign Facebook accounts using AI Pauline Hanson to manipulate Australians](#), *ABC News*, 11 March 2026.

⁷⁹ B Chesney and D Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 2019, Vol 107, 1753-1820, p 1786. (Italics in original. Footnotes omitted). See also: C Vaccari and A Chadwick, [Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news](#), *Social Media and Society*, 2020, Vol 6(1). See also: M Platow et al, [Information resilience: How misinformation undermines social cohesion, trust and democracy](#), 2026, report to the National Science and Technology Council, Australian Government Office of the Chief Scientist, especially Ch 1.4.

⁸⁰ M Pawelec, [Deepfakes and democracy \(theory\): How synthetic audio-visual media for disinformation and hate speech threaten core democratic functions](#), *Digit Soc*, Sep 2022, 8(1), 19.

⁸¹ B Chesney and D Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 2019, Vol 107, 1753-1820, p 1786. (Italics in original. Footnotes omitted). See also: C Vaccari and A Chadwick, [Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news](#), *Social Media and Society*, 2020, Vol 6(1).

⁸² A Ray, [Disinformation, deepfakes and democracies: The need for legislative reform](#), *UNSW Law Journal*, 2021, Vol 44(2), p 1,007.

⁸³ A Ray, [Disinformation, deepfakes and democracies: The need for legislative reform](#), *UNSW Law Journal*, 2021, Vol 44(2), p 1,007.

⁸⁴ J Parker, [A licence for the satirist? The failure of Australian defamation law to protect satire in political media](#) *UNSW Law Journal*, 2021, 21, citing Justice Tony Fitzgerald, 'Telling the truth, laughing' (1999) 92 (August) *Media International Australian incorporating Culture and Policy* 11, 11. Parker notes that satire 'has remained relevant across varying political systems and cultures in the Western World since its origin in Graeco-Roman civilisation'.

5. Law reform in New South Wales

5.1 New electoral law offences

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.⁸⁵ When introducing the bill, 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. We have all seen examples where deepfakes of politicians or other public figures have been used to mislead the public into believing the depicted person said or did something that they did not.⁸⁶

The government noted that similar offences had been introduced in South Australia.⁸⁷

There are 2 new offences in NSW:

1. New 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 simulated person performing an act that the real person depicted did not perform.
2. New 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.⁸⁸

The election period means the period from the date of the issue of the writ for the election until 6 pm on election day.⁸⁹ Electoral material means anything containing 'electoral matter'. 'Electoral matter'

⁸⁵ [Electoral Legislation Amendment \(Elections\) Act 2026](#)

⁸⁶ J Aitchison, [Electoral Legislation Amendment \(Elections\) Bill 2026](#), *NSW Hansard: Legislative Assembly debates*, 17 March 2026.

⁸⁷ J Aitchison, [Electoral Legislation Amendment \(Elections\) Bill 2026](#), *NSW Hansard: Legislative Assembly debates*, 17 March 2026.

⁸⁸ New section 189A(4). The government explained that it 'does not intend to prescribe certain types of excluded electoral material in a regulation at this time. However, given the rate at which AI technologies are developing, we think it is prudent to include that provision to ensure that certain types of digitally generated material may be carved out in the future if required: J Aitchison, [Electoral Legislation Amendment \(Elections\) Bill 2026](#), *NSW Hansard: Legislative Assembly debates*, 25 March 2026.

⁸⁹ [Electoral Act 2017](#), s 4.

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.⁹⁰ A simulated person includes a person depicted in digitally generated electoral material that 'purports to be a depiction of a particular real person'.⁹¹

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.

The maximum penalties are:

- In relation to the first offence – for an individual, 60 penalty units (\$6,600) and/or imprisonment for 6 months; and otherwise 300 penalty units (\$33,000)
- In relation to the second offence – for an individual, 20 penalty units (\$2,200) and/or imprisonment for 6 months; and otherwise 100 penalty units (\$11,000).

The bill passed parliament on 26 March 2026, and these provisions commenced on assent.

5.2 Constitutional issues

The High Court has recognised that the Australian Constitution contains an implied freedom of political communication. Commonwealth or state legislation that breaches the implied freedom can be struck down by the courts. A law will breach the implied freedom if: (1) the law effectively burdens freedom of political communication, and (2) the law does not have a legitimate purpose or the law is not reasonably appropriate and adapted to achieve a legitimate purpose.⁹²

Some NSW electoral laws have been successfully challenged for breaching the implied freedom. For example, in *Unions NSW v New South Wales* the court struck down a provision which made it unlawful for a political donation to be accepted unless the donor was an individual who was enrolled to vote.⁹³ The court concluded this provision was not appropriate and adapted to achieving a legitimate purpose because it was broad and not obviously directed to possible corruption.⁹⁴

Most of the parliamentary debate about the NSW offences on AI-generated electoral material was about the implied freedom of political communication. The Opposition argued that the new offences

⁹⁰ [Electoral Act 2017](#), s 4.

⁹¹ [Electoral Act 2017](#), new sections 189A(4) and 189B(4).

⁹² See *Brown v Tasmania* (2017) 261 CLR 328 and *Clubb v Edwards* (2019) 267 CLR 171.

⁹³ See *Unions NSW v New South Wales* (2013) 252 CLR 530 and L Roth, [The High Court's decision in the electoral funding law case](#), NSW Parliamentary Research Service, e-brief 2/2014, February 2014.

⁹⁴ For an example of a case where the High Court held that laws did not breach the implied freedom, see [LibertyWorks Inc v Commonwealth of Australia](#) [2021] HCA 18.

would breach the implied freedom and moved amendments to create an additional defence of substantial truth or fair comment.⁹⁵ The government did not support the amendments, stating:

Any scheme to regulate electoral material must be carefully limited and proportionate to ensure that it does not suppress robust political debate, satire, parody or legitimate opinion. It must be targeted to ensure that it does not have a chilling effect on lawful political communication. The bill achieves that.⁹⁶

A 2021 journal article discussed the implied freedom of political communication in the context of political deepfake laws generally.⁹⁷ 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.'⁹⁸ He proposed an offence of publishing an altered image if it 'contains a statement regarding electoral matters that is inaccurate or misleading to a material extent.'⁹⁹ 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'.

5.3 Other laws in NSW

As a result of reforms in 2025 there are new offences in the [Crimes Act 1900 \(NSW\)](#) that prohibit the creation and distribution of sexually explicit deepfakes.¹⁰⁰ The Commonwealth [Online Safety Act 2021](#) also provides the eSafety Commissioner with regulatory powers in relation to online image-based abuse. Both laws would apply to sexually explicit deepfakes targeting politicians, but not other types of political deepfakes. The subject of a political deepfake might also be able to pursue a civil action in defamation, passing off, or copyright law. However, Ray explains that there are several limitations associated with these actions in the context of political deepfakes, such as 'the time and costs needed to bring a defamation action.'¹⁰¹

⁹⁵ A Henskens, [Electoral Legislation Amendment \(Elections\) Bill 2026](#), *NSW Hansard: Legislative Assembly debates*, 25 March 2026.

⁹⁶ J Aitchison, [Electoral Legislation Amendment \(Elections\) Bill 2026](#), *NSW Hansard: Legislative Assembly debates*, 25 March 2026.

⁹⁷ A Ray, [Disinformation, deepfakes and democracies: The need for legislative reform](#), *UNSW Law Journal*, 2021, 44(3): 983-1013, p 1005-1007.

⁹⁸ A Ray, [Disinformation, deepfakes and democracies: The need for legislative reform](#), *UNSW Law Journal*, 2021, 44(3), p 1005-1007.

⁹⁹ A Ray, [Disinformation, deepfakes and democracies: The need for legislative reform](#), *UNSW Law Journal*, 2021, 44(3): p 1011.

¹⁰⁰ [Crimes Act 1900 \(NSW\)](#), section 91PA introduced by [Crimes Amendment \(Intimate Image and Audio Material\) Act 2025](#). For background see T Gotsis, [Sexually explicit deepfakes and the criminal law in NSW](#), NSW Parliamentary Research Service, April 2025. See also [Criminal Code Act 1995 \(CTH\)](#), sections 474.17A and 474.17AA.

¹⁰¹ A Ray, [Disinformation, Deepfakes and democracies: The need for legislative reform](#), *UNSW Law Journal*, 2021, 44(3): 983-1013, p 1000. See also W Potter, [AI deepfakes threaten democracy and people's identities. 'Personality rights' could help](#), *The Conversation*, 5 March 2025.

6. Law reform in other jurisdictions

6.1 Australia

6.1.1 South Australian electoral law reform

South Australia is the only other Australian jurisdiction that has reformed its electoral laws to address political deepfakes. These reforms were enacted to supplement existing truth in political advertising laws.¹⁰² 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.¹⁰³ An 'electoral advertisement' is an advertisement containing matter calculated to affect the result of an election. There are similar defences to those in NSW. The maximum penalty for the offence is a fine of up to \$5,000 for an individual (\$10,000 for a body corporate).

6.1.2 National law reform proposals

An October 2024 Senate Select Committee inquiry on AI examined the potential impacts of AI on democracy, including the use of deepfakes.¹⁰⁴ One of its recommendations was that:

...the Australian Government undertake a thorough review of potential regulatory responses to AI-generated political or electoral deepfake content, including mandatory codes applying to the developers of AI models and publishers including social media platforms, and prohibitions on the production or dissemination of political deepfake content during election periods, for legislative response prior to the election of the 49th Parliament of Australia.¹⁰⁵

In November 2024 the Australian Government introduced a bill to enact truth in political advertising laws, which included specific provisions on deepfakes.¹⁰⁶ 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.¹⁰⁷ Authorisable electoral matter mainly included advertisements in which all or part of the distribution or production was paid for.¹⁰⁸ Electoral matter would not include a communication for a dominant purpose that is a satirical, academic, educative or artistic purpose.¹⁰⁹ Breaches of the

¹⁰² South Australia and the ACT are the only Australian jurisdictions that are considered to have truth in political advertising laws. See Yee-Fui Ng, [Truth in political advertising laws: design, operation and effectiveness and recommendations for reform: Final report](#), report prepared for the Susan McKinnon Foundation, December 2024.

¹⁰³ [Electoral \(Miscellaneous\) Amendment Act 2024](#). The reforms followed the release in August 2024 of a [discussion paper](#) on the regulation of deepfakes in various contexts.

¹⁰⁴ Senate Select Committee on Adopting Artificial Intelligence, [Select Committee on Adopting Artificial Intelligence – Interim Report](#), Parliament of Australia, October 2024, Ch 2.

¹⁰⁵ Senate Select Committee on Adopting Artificial Intelligence, [Select Committee on Adopting Artificial Intelligence – Interim Report](#), Parliament of Australia, October 2024, p 33 (Rec 2).

¹⁰⁶ [Electoral Legislation Amendment \(Electoral Communications\) Bill 2024](#)

¹⁰⁷ See proposed new sections 321M and 321MA.

¹⁰⁸ See proposed new section 321JA(2).

¹⁰⁹ See note to proposed section 321JA(2) referring to section 4AA(5).

deepfake provisions would attract a civil penalty of up to 1,000 penalty units (fine of \$330,000). This was substantially higher than the penalties now in place in NSW and South Australia.

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.¹¹⁰ The scheme includes civil penalties (500 penalty units), and powers for the eSafety Commissioner to issue removal notices to social media service providers and hosting service providers. As at 21 April 2026 the bill had not been debated.

6.2 International

6.2.1 Electoral law reforms

In the United States, 26 states have enacted political deepfake laws. In summary:

States have generally taken two approaches: prohibitions and disclosures. Two states—Minnesota and Texas — prohibit the publication of political deepfakes a certain number of days prior to an election. The other 24 states require disclosures on the media, like those required for who paid for a political ad, stating if it contains a deepfake...¹¹¹

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.¹¹² For example, in April 2025 X Corp (formerly Twitter) challenged the Minnesota laws.¹¹³ As at 21 April 2026 this case has not been decided.

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.¹¹⁴ 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.

Laws banning political deepfakes have also been enacted in South Korea¹¹⁵ and Singapore.¹¹⁶ The Singapore laws empower the Returning Officer to issue a written direction to require a social media

¹¹⁰ [Online Safety and Other Legislation Amendment \(My Face, My Rights\) Bill 2025](#). See also D Pocock, [New Bill to protect identity in deepfake future](#) [media release], 24 November 2025. Note also that Zali Steggall had previously introduced the [Commonwealth Electoral Amendment \(Stop the Lies\) Bill 2021](#).

¹¹¹ National Conference of State Legislatures, [Deepfakes in elections and campaigns](#), updated 15 December 2025. For proposed laws at the federal level, see R Sam Garrett, [Artificial Intelligence \(AI\) and Campaign Finance Policy: Recent Developments](#), Congressional Research Service, updated 25 September 2024.

¹¹² Y Kim, [The legal gray zone of deepfake political speech](#), *Cornell Journal of Law. & Public Policy: The Issue Spotter*, 24 October 2025.

¹¹³ S Karnowski, [Elon Musk's X sues to overturn Minnesota political deepfakes ban](#), AP, 26 April 2025.

¹¹⁴ The [Election Financing Amendment and Elections Amendment Act](#), which added a new section 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.

¹¹⁵ See [Public Official Election Act](#), 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

¹¹⁶ See [Parliamentary Elections Act 1954](#), s 61MA, which was inserted by the [Elections \(Integrity of Online Advertising\) \(Amendment\) Act 2024](#).

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.¹¹⁷

6.2.2 Other law reforms

European Union – laws on digital services and AI

The European Union's [Digital Services Act](#) (2022) 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.¹¹⁸ Such mitigation measures may include ensuring that a generated or manipulated image, audio or video that appreciably resembles existing persons is distinguishable through prominent markings.¹¹⁹ In 2024 the European Commission published guidelines under the Act on the mitigation of systemic risks for electoral processes.¹²⁰

The European Union's [Artificial Intelligence Act](#) (2024) requires developers of AI systems generating synthetic content to ensure that the outputs are marked in a machine-readable format and detectable as artificially generated or manipulated.¹²¹ It also obliges deployers of an AI system that generates a deepfake to disclose that it has been artificially generated or manipulated.¹²² A deployer is a user of an AI system except where it is used 'in the course of a personal non-professional activity'.¹²³ It has been noted that this exception might reduce the Act's effectiveness in tackling deepfakes.¹²⁴

Denmark – changes to copyright laws to address deepfakes

In 2025 the Danish Government proposed amendments to the Copyright Act to address deepfakes in general.¹²⁵ The changes are designed to protect the public against 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. The laws are expected to come into force in July 2026.

¹¹⁷ [Parliamentary Elections Act 1954](#), s 61N.

¹¹⁸ See Article 35. Very large online platforms and online search engines are defined in Article 33.

¹¹⁹ Article 35(1)(k)

¹²⁰ 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](#).

¹²¹ See Article 50(2).

¹²² See Article 50(4).

¹²³ See Article 3(4).

¹²⁴ M Labuz, [Deep fakes and the Artificial Intelligence Act—An important signal or a missed opportunity?](#) *Policy & Internet*, 2024, 16(4):783-800, p 794.

¹²⁵ Library of Congress, [Denmark: political parties agree to protect Danes against deepfakes](#), Global Legal Monitor, 5 August 2025; and S Karttunen, [The Danish approach to copyright and deepfakes: A model for the EU?](#), European Parliamentary Research Service, January 2026.

7. Other responses to political deepfakes

7.1 Deepfake counter measures

There are existing and emerging deepfake counter measures which could help mitigate the potential adverse effects of political deepfakes. They include:

- Invisible watermarks
- Deepfake detection
- Prompt removal of deepfakes from social media platforms
- Journalistic fact checking and reporting of media statements
- Public education and awareness.

7.1.1 Invisible watermarks

Watermarking has been described as an ‘umbrella term encompassing several different approaches to embedding patterns in digital media’.¹²⁶ Essentially, it is ‘the process of concealing or embedding data behind an image or video that is invisible to the naked human eye’.¹²⁷

If invisible watermarks were automatically included in all digital content that was created by AI platforms, they would provide an ‘authentication trail’ that computers could use to identify that particular content was inauthentic because it was generated by AI.¹²⁸

Several challenges to the effective deployment of watermarks have been identified. In particular, the technology is in its ‘infancy’. Watermarks have developed to the point where they are ‘robust to erasure and forgery’, but they ‘are not foolproof ... [because] a motivated actor can degrade watermarks in AI-generated content’.¹²⁹ Further, the co-operation and/or regulation of AI developers would also be required.¹³⁰

¹²⁶ S Srinivasan, [Detecting AI fingerprints: A guide to watermarking and beyond](#), *Brookings*, 4 January 2024.

¹²⁷ P 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 2-9. See also Resemble AI, [What is AI watermarking and why it matters in 2026?](#), n.d., accessed 9 April 2026.

¹²⁸ That deepfake technology would create a need for authentication trails was foreshadowed by Chesney and Citron in 2019 when they said: ‘... [in] a world in which it is cheap and easy to portray people as having done or said things they did not say or do ... a person who cannot credibly demonstrate their real location, words, and deeds at a given moment will be at greater risk than those who can. Credible alibis will become increasingly valuable ... We predict the development of a profitable new service: immutable life logs or authentication trails that make it possible for a victim of a deep fake to produce a certified alibi credibly proving that he or she did not do or say the thing depicted ...’: B Chesney and D Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 2019, Vol 107, 1753-1820, p 1,814.

¹²⁹ S Srinivasan, [Detecting AI fingerprints: A guide to watermarking and beyond](#), *Brookings*, 4 January 2024.

¹³⁰ As discussed earlier (at 6.2.2), Article 50(2) of the European Union’s [Artificial Intelligence Act](#) (2024) requires developers of AI systems generating synthetic content to ensure that the outputs are marked in a machine-readable format and detectable as artificially generated or manipulated. Similar requirements could specifically require that AI generated content be embedded with watermarks. See also: J Evans, [Tech companies advised to label and ‘watermark’ AI-generated content](#), *ABC news*, 1 December 2025.

7.1.2 Deepfake detection

Deepfake detection technologies are designed to reveal the origins of digital content, regardless of whether information has been attached to it or embedded in it.¹³¹

A 2026 report by the UK Government notes that deepfake detection technologies are ‘critical for safeguarding democratic process’ but the ‘development of deepfake detection remains in its early stages, and the market itself remains nascent’.¹³² It also notes that the accuracy of deepfake detection technology is hampered by limited access to high quality training datasets, inconsistencies in evaluation metrics and technology developers claiming high accuracy without independent evaluation.¹³³ It adds that a key challenge faced by the deepfake detection industry is to establish confidence in the accuracy of the technology because ‘users generally lack confidence in the ability of deepfake detection tools to distinguish real from manipulated content.’¹³⁴

7.1.3 Prompt removal of deepfakes from social media platforms

The experience of Catherine Connolly in the 2025 Irish presidential election suggests that the adverse effects of political deepfakes may be reduced by their prompt removal from social media platforms.¹³⁵ In Australia, the e-Safety Commissioner has takedown or removal powers relating to cyber abuse (such as threats of violence), image-based abuse (non-consensual sexually explicit material) or illegal and restricted online content (such as extreme violence or child abuse material).¹³⁶ In other cases a decision to remove a political deepfake rests with the social media platform.

7.1.4 The role of the media: fact checking and reporting of media statements

As discussed earlier (at 4.2.1), a 2025 experimental study showed that the ‘substantial reputational damage’ caused by particularly toxic political deepfake videos could be substantially reduced or even eliminated by journalistic fact checking.¹³⁷ Such findings suggest that 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.¹³⁸ The BBC is developing its own in-house deepfake detection tools in an effort to uphold the accuracy and objectivity of its journalism.¹³⁹ Other media

¹³¹ United Kingdom Government, [Deepfake detection technology](#), 26 March 2026. See further: F Romero-Moreno, [Deepfake detection in generative AI: A legal framework proposal to protect human rights](#), *Computer Law & Security Review*, September 2025, Vol 58.

¹³² United Kingdom Government, [Deepfake detection technology](#), 26 March 2026.

¹³³ United Kingdom Government, [Deepfake detection technology](#), 26 March 2026. See also: CSIRO, [Research reveals ‘major vulnerabilities’ in deepfake detectors](#), 13 March 2025, J Bowler, [Deepfake detectors struggle to tell real from fake using real world data](#), *ABC News*, 17 March 2025 and [No tech fix yet: Deepfakes are outpacing detection systems](#), *Devdiscourse*, 8 April 2026.

¹³⁴ United Kingdom Government, [Deepfake detection technology](#), 26 March 2026 (footnotes omitted).

¹³⁵ As discussed above at 4.1.

¹³⁶ See eSafety Commissioner, [Summary table of what you can report and how](#), 6 September 2026. See also, eSafety Commissioner: [Adult cyber abuse](#), 19 February 2025; [Report image-based abuse](#), 2 January 2026; and [Illegal and restricted online content](#), 4 December 2025.

¹³⁷ V Dan, [Deepfakes as a democratic threat: Experimental evidence shows noxious effects that are reducible through journalistic fact checks](#), *The International Journal of Press/Politics*, 11 February 2025.

¹³⁸ P Raemy et al, [Deepfakes and journalism: Normative considerations and implications](#), *Journalism studies*, 2025, vol 26(14), 1742-1744.

¹³⁹ W Bayliss, [Deepfake detection for journalism: How we’re tackling manipulated media](#), *BBC*, 5 November 2025.

outlets are using third-party fact checking tools and services to verify the authenticity of information.¹⁴⁰

7.1.5 Public education and awareness

A 2025 study examined the effectiveness of deepfake public education and awareness. It exposed one-third of participants to text-based information about deepfakes ('passive inoculation'), another third of participants to an interactive game that challenged them to identify deepfakes ('active inoculation'), while the remaining participants were not exposed to any deepfake education and awareness ('no inoculation'). All participants were then randomly shown one of 2 different deepfakes. The study found that 'both types of inoculation were effective in reducing the credibility participants gave to the deepfakes, while also increasing people's awareness and intention to learn more about them.'¹⁴¹

The Senate Committee's 2024 report on the impact of AI on democracy recommended that the Australian Government examine mechanisms to improve AI literacy for all Australians, particularly in an electoral context.¹⁴² The UK Government has said that public education and awareness is a 'useful tool for increasing awareness and recognition of AI-generated disinformation ... and build[ing] people's resilience to disinformation'.¹⁴³ It added that 'citizens may welcome tools that help them to report deepfakes – not just for their own media literacy, but to contribute to research and improve social outcomes.' Some stakeholders have suggested, however, that 'technological advances are becoming so sophisticated it is unreasonable to expect consumers and citizens to be able to "spot" deceptive imagery and voices'.¹⁴⁴

7.2 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*.¹⁴⁵ 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]

¹⁴⁰ A Yazdinejad and J Kong, [Battling deepfakes: How AI threatens democracy and what we can do about it](#), *The Conversation*, 21 August 2025 and J Adwell, [Journalism in the age of AI fact-checking: Navigating truth, tools and trust](#), *The Daily Mesh*, 28 October 2025.

¹⁴¹ D Zhang, ['Inoculation' helps people spot political deepfakes, study finds](#), *The Conversation*, 5 February 2026. See further: B Zhang, S Kim and A Scott, [Immunizing the Public Against AI-Generated Disinformation: Testing the effects of inoculation mode and issue attitude on inoculation likelihood of political deepfakes](#), *Journalism and Mass Communication Quarterly*, 2025, Vol 102(4) (links to abstract only).

¹⁴² Senate Select Committee on Adopting Artificial Intelligence, [Select Committee on Adopting Artificial Intelligence – Interim Report](#), Parliament of Australia, October 2024, p 35, Rec 5.

¹⁴³ United Kingdom Government, [Deepfakes and media literacy](#), 27 May 2025, accessed 13 April 2026.

¹⁴⁴ Inter-Parliamentary Union, [Dangers of deepfakes for parliamentarians](#), 19 February 2024, quoting Sam Gregory, Executive Director of WITNESS, an international nonprofit organization that assists people use technology to protect and defend human rights.

¹⁴⁵ Munich Security Conference, [A tech accord to combat deceptive use of AI in 2024 elections](#), February 2024.

shared evidence of their actions, inconsistent follow-through, vague reporting, and a lack of independent verification made it challenging to assess real progress.’¹⁴⁶

The Australian Communications and Media Authority (ACMA) monitors the activities of digital platforms under the *Australian Code of Practice on Disinformation and Misinformation*, which is not legally binding.¹⁴⁷ ACMA’s latest report was published in August 2025 but only captures activities in 2024. The report noted that generally signatories ‘are deploying systems and processes to detect and label AI-generated disinformation and misinformation’.¹⁴⁸ The next report to be published later in 2026 will cover the period of the 2025 Australian federal election.

7.3 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.¹⁴⁹ The NSW Electoral Commission also had a ‘Stop and Consider’ campaign for the 2023 state election.¹⁵⁰ Some international electoral authorities are also exploring 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.¹⁵¹ The article also noted that the Scottish Electoral Commission had asked the government to consider giving it legally enforceable ‘takedown’ powers to require social media platforms to remove deepfake material.¹⁵²

¹⁴⁶ A Ahmed et al, [Tech companies pledged to protect elections from AI – here’s how they did](#), Brennan Center for Justice, Policy Brief, 13 February 2025.

¹⁴⁷ Digital Industry Group Inc (DIGI), [Australian code of practice on disinformation and misinformation](#), July 2024.

¹⁴⁸ Australian Communications and Media Authority, [Digital platforms’ efforts under voluntary arrangements to combat disinformation and misinformation: Fourth report to government](#), August 2025, p 2.

¹⁴⁹ Australian Electoral Commission, [AI & elections](#), accessed 23 March 2026.

¹⁵⁰ NSW Electoral Commission, [Commission campaign to tackle election disinformation](#) [media release], 21 February 2023.

¹⁵¹ S Carrell, [Software tackling deepfakes to be piloted for Scottish and Welsh elections](#), *The Guardian*, 9 January 2026.

¹⁵² S Carrell, [Software tackling deepfakes to be piloted for Scottish and Welsh elections](#), *The Guardian*, 9 January 2026.

8. Conclusion

Political deepfakes pose a new challenge to democratic institutions. They have the potential to influence elections, harm politicians, and erode trust in political communication. 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 the publication of political deepfakes that could deceive voters. In addition to laws, combatting political deepfakes is likely to require a range of actions, including technological measures and public awareness 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.

**Political deepfakes and the new laws in
NSW**

Lenny Roth and Tom Gotsis

Research Paper No. 2026-02

ISSN 2653-8318

© 2026 Except to the extent of the uses permitted under the Copyright Act 1968, no part of this document may be reproduced or transmitted in any form or by any means including information storage and retrieval systems, without the prior consent from the Senior Manager, NSW Parliamentary Research Service, other than by members of the New South Wales Parliament in the course of their official duties.

Disclaimer: Any advice on legislation or legal policy issues contained in this paper is provided for use in parliamentary debate and for related parliamentary purposes. This paper is not professional legal opinion.

The NSW Parliamentary Research Service provides impartial research, data and analysis services for members of the NSW Parliament.

parliament.nsw.gov.au

Media inquiries should be directed to:
media@parliament.nsw.gov.au

The Parliament of New South Wales acknowledges and respects the traditional lands of all Aboriginal people and pays respects to all Elders past and present. We acknowledge the Gadigal people as the traditional custodians of the land on which the Parliament of New South Wales stands.

This image comes from 'Our Colours of Country', which was created for the Parliament of NSW by Wallula Bethell (Munro) a Gumbaynggirr/Gamilaroi artist born and raised in Tamworth who has spent time living on Dunghutti Country and is currently living in Western Sydney on Darug Country with her husband and son.

