# **Artificial Intelligence**

This chapter explores the interaction between artificial intelligence (AI) and the internet, with a focus on the current growth in AI, the regulation of AI and associated challenges, and how AI can be used as a regulatory tool.

## **Brief history of AI**

The first rudimentary artificial intelligence was conceptualised in 1943 in the development of an artificial neuron.[[1]](#footnote-1) The term ‘artificial intelligence’ itself, however, was not used until 1956 when John McCarthy used the terms to describe ‘the science and engineering of making machines intelligent’.[[2]](#footnote-2) Today the term artificial intelligence refers to the ability for computer systems to assess, use and adapt available information to make informed decisions and recommendations about existing circumstances, much like a human would using their own intelligence.[[3]](#footnote-3)

## **Regulation of AI**

The use of AI is rapidly increasing, permeating industries and communities across the globe. The value of AI globally is estimated to reach $282.80 billion by the end of 2024, and projected to grow by an additional 29.29% by 2030, resulting in a market value of 1,321.00 billion.[[4]](#footnote-4)

Due to its rapid growth and development, regulatory mechanisms have struggled to keep up. However, the implementation of AI regulations has seen a sharp rise in recent years. For example, in the United States, in 2016 there was a single AI related regulation, whereas in 2023 this had increased to 25 regulations, with a 56% increase in that year alone.[[5]](#footnote-5)

1. **Australia**

Australian lawmakers have started to regulate the use of AI. The Online Safety (Basic Online Safety Expectations) Determination 2022,[[6]](#footnote-6) made under section 45 of the Online Safety Act 2022,[[7]](#footnote-7) is an attempt to ensure the safety of internet users and reduce the risk of misuse of AI online. The Determination requires that internet service providers take reasonable steps to ensure that their delivery of AI to the consumer base has safety at the forefront of its design implementation and maintenance. Reasonable steps include undertaking safety assessments, providing educational tools for users, monitoring the data used as training material for the AI systems, and implementing ways of detecting harmful content. The Determination also requires that providers are proactive in their approaches to reducing the risk of their AI being used to create harmful content.[[8]](#footnote-8)

Other Australian state and national legislation such as the *Privacy Act,*[[9]](#footnote-9) the Privacy and Personal Information Protection Act 1998[[10]](#footnote-10) (NSW) and the Privacy Legislation Amendment (Enforcement and Other Measures) Act 2022[[11]](#footnote-11) can assist in regulating certain elements of AI use on the internet, but there is not yet any specific legislation that directly regulates the use of AI online. However, the Australian government is currently engaged in consultation around the safe use of AI in Australia and other policy activity. Notably:

* June 2023: a Discussion paper was released on the safe use of AI for public comment.[[12]](#footnote-12) Topics include the opportunities and challenges of AI, and strategies for managing the risks posed by AI.
* September 2023: the [AI in Government Taskforce](https://www.dta.gov.au/blogs/ai-government-taskforce-examining-use-and-governance-ai-aps) was set up
* January 2024: commentary was released by the government acknowledging the challenges presented by AI and detailing the potential mechanisms which may be required for the use of AI.[[13]](#footnote-13)
* February 2024: the [Artificial Intelligence Expert Group](https://www.industry.gov.au/science-technology-and-innovation/technology/artificial-intelligence#meet-the-ai-advisory-expert-group-4) was set up.
* June 2024: “The National framework for the assurance of artificial intelligence in government” was released. This framework deals with the government’s use of AI.

1. **Regulatory challenges**
2. **Misinformation and Disinformation**

**See** [**an explanation of misinformation**](https://www.acma.gov.au/online-misinformation) **by the ACMA.**

Artificial Intelligence (‘AI’) has contributed to the ongoing challenge of regulating and controlling the spread of misinformation and disinformation. Misinformation is ‘false, misleading or deceptive information that can cause harm’. Disinformation is misinformation that is deliberately spread to cause confusion and undermine trust in governments or institutions.[[14]](#footnote-14) Algorithms and ‘bots’ are becoming some of the strongest spreaders of false, unreliable and misleading information online. ‘Bots’ are computer algorithms generated by AI that automatically produce content and interact with humans on social media platforms.

The spread of misinformation and disinformation online has been linked to propaganda and the proliferation of abuse and targeted attacks, and harm in emergency situations as civilians are unable to obtain the correct information from reliable sources about how they should ensure their own safety. A 2023 [Forbes report](https://www.forbes.com/advisor/business/artificial-intelligence-consumer-sentiment/#:~:text=According%20to%20the%20survey%20data%2C%20a%20combined%2076%25,43%25%20being%20very%20concerned%20and%2033%25%20somewhat%20concerned.) indicated that 76% of consumers were worried about misinformation provided by AI.

***Regulation***

Misinformation and disinformation are regulated in Australia through a voluntary code of practice. The *Australian Code of Practice on Disinformation and Misinformation* (The Code) was released in Australia by The Digital Industry Group (DIGI) in February 2021. DIGI is a not-for-profit industry association tasked with administering the Code. The objective of the Code is to combat false material being released on digital platforms by setting a standard of practice to which signatories are required to comply with. Eight technology companies have opted into commitments under the Code, however according to provision 7.1 they are only required to comply with their selected commitments. Provision 7.2 also recognises that companies may withdraw from the Code by notifying DIGI. An independent Complaints Committee resolves complaints regarding Signatories compliance with their commitments under the Code and the public has access to complaints that are made via a complaints portal on DIGI’s website.[[15]](#footnote-15) The Australian Communications and Media Authority (ACMA) also has oversight over the code and reports on the adequacy of platforms measures to implement their commitments. These reports are then publicly available.

***Online Safety Act***

The *Online Safety Act 2021* (Cth) does not directly regulate the spread of misinformation and disinformation. However, the Commissioner has the power to require providers to report on the extent to which they are complying with expectations under BOSE. Importantly, however, failure to comply with the expectations listed in BOSE will not lead to legal penalties as they are not enforceable by proceedings in a court.[[16]](#footnote-16) Inclusive within the ‘Core Expectations’ under BOSE, providers are required to take ‘reasonable steps’ to ensure the safety of their end-users and to prevent ‘harmful material’ being released on their sites.[[17]](#footnote-17) Under the first determination of BOSE in 2022 the Minister for Communications set out expectations that providers would take reasonable steps to minimise the extent to which AI and anonymous accounts would produce harmful material on their sites.[[18]](#footnote-18) ‘Harmful material’ is not defined anywhere in the Act, however it is considered a ‘reasonable step’ by the provider to request a consultation with the eSafety Commissioner in making determinations about what may be ‘harmful’.

***Recent Developments***

The Australian Government appears to be interested in introducing legislation to combat the spread of misinformation and disinformation on digital platforms. A senate inquiry was conducted in 2023 by the Economics References Committee who reported on the ‘Influence of international digital platforms’. This inquiry received several submissions from organisations such as the Human Rights Law Centre noting concerns about the rise of disinformation and misinformation online and the failure of any existing effective enforcement mechanism combatting it. An exposure draft of the *Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill* 2023 was released for public feedback on 25 June 2023. This legislation would afford ACMA new powers to hold digital platforms to account and strengthen and support the Code to extend to non-signatories. The Government, however, has not yet announced a timeline for introduction of this Bill to parliament and there is considerable pushback from organisations noting concerns for the restraint it may impose on freedom of expression.

**See an outline of** [**changes to expect in 2024**](https://www.aph.gov.au/About_Parliament/Parliamentary_departments/Parliamentary_Library/Research/FlagPost/2024/March/Media_regulation_2024) **by Nell Fraser posted to the Australian Parliament’s website.**

**ii) DEEPFAKES COULD GO HERE? SEE 007 and 008**

## **AI as a Regulatory Tool**

AI may be used to assist in regulating the internet. There are many examples of this, including

1. Fraud detection. For example, banks will use AI in real time to assess patterns of behaviour to determine whether fraudulent activity is taking place.[[19]](#footnote-19)
2. Spam filtering. For example AI will use learned algorithms to analyse massive amounts of data to identify characteristics, patterns and anomalies which may indicate spam.[[20]](#footnote-20)
3. Behavioural Patterns. For example, PayPal uses AI to monitor behavioural patterns of its users to identify potential fraudulent behaviour. If changes in spending patterns such as a large or out of character transaction is made, the transaction can be frozen pending authorisation.[[21]](#footnote-21)
4. Content regulation is another area.

**Case study: Algorithmic moderation as content regulation**

“Algorithmic moderation” refers to the use of automated systems, typically powered by machine learning algorithms and artificial intelligence (**AI**) to monitor, evaluate and manage online content. These systems are designed to detect and take action against content that violates platform policies, such as hate speech, misinformation, or explicit material. Unlike human moderators, algorithmic moderation can process vast amounts of content in real-time, making it an essential tool for large-scale platforms like social media networks.

The concerns surrounding algorithmic moderation stem from its potential for errors and biases, which can result in the wrongful removal of legitimate content or the failure to detect harmful material. The implications of these errors are amplified by the vast reach of the internet, where decisions made by algorithms can impact millions of users in real time.

Whilst algorithmic moderation as a form of content moderation may be effective in removing illegal content, it has systemically struggled in removing harmful content. Like with human moderation, it can be difficult to differentiate between what is “harmful” and what is merely a non-mainstream opinion.

1. Warren S. McCulloch and Walter H Pitts, ‘A Logical Calculus if The Idea Immanent in Nervous Activity’ (Research Paper, Vol 5, Bulletin of Mathematical Biophysics, 1943) 115-133. [↑](#footnote-ref-1)
2. John McCarthy. “What is artificial intelligence” (Article, Computer Science Department, Stanford University, 12 November 2004). [↑](#footnote-ref-2)
3. Iberdrola, ‘Artificial Intelligence: birth, applications and future trends’, *History of Artificial Intelligence (Blog Post).* [↑](#footnote-ref-3)
4. Statista, ‘Artificial Intelligence – Worldwide’, *Market Insights* (Web Page, 2024) https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide?currency=AUD. [↑](#footnote-ref-4)
5. Stanford University, *Artificial Intelligence Index Report 2024* (2024). [↑](#footnote-ref-5)
6. Online Safety (Basic Online Safety Expectations) Determination 2022 (Cth) (“The Determination”). [↑](#footnote-ref-6)
7. *Online Safety Act 2021* (Cth). [↑](#footnote-ref-7)
8. n 6. [↑](#footnote-ref-8)
9. *Privacy Act 1988* (Cth). [↑](#footnote-ref-9)
10. *Privacy and Personal Information Protection Act* 1998 (NSW). [↑](#footnote-ref-10)
11. *Privacy Legislation Amendment (Enforcement and Other Measures) Act 2022* (Cth). [↑](#footnote-ref-11)
12. Australian Government Department of Industry, Science and Resources, *Safe and Responsible AI in Australia* (Discussion Paper, June 2023) <https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public\_assets/Safe-and-responsible-AI-in-Australia-discussion-paper.pdf>. [↑](#footnote-ref-12)
13. Julian Lincoln, Susannah Wilkinson and Alex Lundie, “Australia Government announces mandatory regulations for high-risk AI’ (Article, Insight Australia, 18 January 2024). [↑](#footnote-ref-13)
14. ACMA, Online Misinformation, [www.acma.gov.au/online-misinformation](http://www.acma.gov.au/online-misinformation) (accessed 01 September 2024) [↑](#footnote-ref-14)
15. *Australian Code of Practice on Disinformation and Misinformation* s 7.5. [↑](#footnote-ref-15)
16. *Online Safety Act 2021* (Cth) s 45(4) [↑](#footnote-ref-16)
17. *Online Safety Act 2021*(Cth) s 46(1)(a)(b) [↑](#footnote-ref-17)
18. *Online Safety (Basic Online Safety Expectations) Determination* 2022 s8A, s9 [↑](#footnote-ref-18)
19. Ravi Sandepudi, ‘The Banker’s Guide: Using AI for Fraud Detection (Effective, 11 March 2024). [↑](#footnote-ref-19)
20. David Emelianocm, ‘Advanced Spam Filtering AI, *Trimbox* (Blog Post, 21 November 2023). [↑](#footnote-ref-20)
21. Ashtynn Baltimore, ‘Is AI changing customer expectations?’ (2024) *PayPal Braintree Product Team.* [↑](#footnote-ref-21)