‘Real’ Laws for Artificial Intelligence

An introductory guide to AI regulation

February 2024

www.pwc.com.au

Introduction

Artificial Intelligence (AI) is already changing the world – whether that is for better or for worse will depend on how effectively and responsibly its use and development is regulated and governed

ChatGPT, AutoGPT, Llama, Generative AI… In what seems like a whirlwind couple of months, these once (largely) unknown terms have become mainstream in everyday conversations, and this is only the beginning. With the applications and use cases for generative AI only growing, it is critical that governments and businesses alike consider what this means for organisations, people and society and implement appropriate measures to address this.

The key challenge for governments is how do they regulate this new technology in a manner that facilitates innovation and improvement whilst ensuring safety, transparency and reliability. How do we make AI work for society in a positive way? Or, at least, in manner that addresses potential negative consequences.

In this article, we consider AI in the current legal landscape in Australia, key legal concerns with AI and how governments around the world are moving to regulate it. We also look at how entities can proactively implement governance around AI to get the most out of AI in their business in a responsible way and ensure they get a head start on legal compliance.

Key takeaways from our article include

- AI, and in particular, generative AI, is different from traditional software tools. Factors such as complexity of algorithms, lack of transparency, hyper-scalability and self evolution combine to aggravate risk of the harm that it could cause to people, organisations and society.
- There are already a number of existing laws that intersect with AI in Australia. Whilst they still apply and must be considered, many of these legal regimes have not considered the complexities of AI.
- Governments around the world are moving to introduce specific regulation to address the development and use of AI with a range of strategies.
- Organisations should be looking to get ahead of the curve and implement AI governance and risk management processes and procedures that reflect best practice (e.g. risk-based approaches).

Contents

01

The case for further regulation

1.1
What is AI and the case for further regulation

Despite the media frenzy surrounding OpenAI's release of ChatGPT and the subsequent explosion of generative AI products, AI has already been among us for decades. For example, machine learning is used every day by Google to power their algorithms to provide more relevant information through the Google search engine. Another example is where AI has enabled computers to communicate with people through voice control devices, such as Amazon Echo. AI has also been used in medicine to study what makes humans healthy and how cancer spreads, by companies like Microsoft[1] and Alphabet,[2] and scientists have used AI to analyse light signals to produce images of black holes. Governments too have implemented AI, for example the Victorian State Government use AI sensors to incorporate safety metrics into traffic signals in real time, and the Victorian Police use automatic number plate recognition technologies in its surveillance of the public.

AI is unlike other software tools. AI technologies pose a range of unique challenges that aggravate the risk of harm to people, organisations and society. These include:

- The **hyper-scalability** of AI-enabled automation, resulting in widespread adverse impacts in the event of faults or errors in the ways AI systems make predictions or decisions.
- Difficulty interrogating and explaining the outputs of **exceptionally complex algorithmic models**, which are not explicitly programmed by human developers.
- Absence of clear principles for the appropriate assignment of **accountability, ownership and liability** for the outputs of AI models, especially where AI systems are created through complex and interdependent value chains.
- **Adaptability and self change** (i.e. 'learning') for AI models means that a 'set and forget' approach like traditional software products doesn't work.

There is no doubt that the power of AI unlocks significant efficiencies and opportunities, with an almost unlimited scope of application. However, with these opportunities comes a plethora of new risks and potential harms that AI presents to organisations, people and society that must be considered. Some examples of these are:

Privacy and data issues

One of the very first concerns that AI developers, users and operators alike must deal with is the privacy issues raised by the development and use of AI, particularly in the area of consent and the incorporation of personal information in AI modelling and inputs.

Bias and discrimination

AI can be swayed by incidental bias and discrimination. There are inherent issues that are built into the input data sets used to train AI algorithms that can give rise to perverse and unwanted outcomes.

Accountability and transparency

The reality of how AI systems are designed (essentially it is a 'black box') and implemented means that the existing standards of accountability, responsibility and liability must be shifted.

Economic and social disruptions

AI will undoubtedly have transformative effects on our economy – even computers, when first introduced, significantly changed many the employment landscape. There are economic and social considerations that must be managed through appropriate AI regulation.

These risks and harms, among others, are the reason why governments around the world are scrambling to establish appropriate controls and guidelines around the use and development of AI enabled tools. An open letter dated 22 March 2023 requesting the immediate moratorium on the training of certain AI systems due to the potential dangers of uncontrolled AI development was signed by more than 1,100 people in a single week, including Apple co-founder Steve Wozniak and technology billionaire Elon Musk. The letter suggested that AI developers should work with policymakers and governments to jointly develop robust AI governance systems and protocols for advanced AI design and development.[3]

In truth, regulating AI is difficult due to the need to strike the balance between allowing for innovation and evolution, whilst ensuring AI is used in a safe and responsible manner. One thing is clear – allowing the unfettered development of AI may be untenable and will likely not be permitted to occur in most, if not, all of world's major economies.

1.2 Defining AI

But what is AI? It seems simple, but it is anything but that.

There are many evolving definitions of AI that have been proposed, but the crux of it usually comes down to a machine or program's ability to imitate processes associated with the intelligent human mind, including learning, reasoning, adapting and self-correction. The definition used to regulate AI, however, will need to be specific, address the targeted AI risks and challenges, and encapsulate how AI is used and applied. The best outcome will include a globally harmonised definition.

Each different type of AI system and the use case that it is applied to carries its own unique risks, and there is much uncertainty of what could still be invented. It is also crucial to reach a correct balance between regulating to protect against harms and encouraging innovation and improvement. The truth is that a clear, sensible and appropriate regime will encourage and accelerate the development of AI through the certainty it provides to developers and users.

A few definitions that have been adopted globally are extracted below:

Example definitions of AI currently used

The AI Act (Council of the European Union)

'artificial intelligence system' (AI system) means a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts.

The AI Act (European Parliament)

'artificial intelligence system' (AI system) means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments.

Organisation for Economic Co-operation and Development (OECD)

An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

Commonwealth Scientific and Industrial Research Organisation (CSIRO)

A collection of interrelated technologies used to solve problems autonomously and perform tasks to achieve defined objectives without explicit guidance from a human being.

02

AI and Australian law

2.0
An overview of the legal issues arising with AI

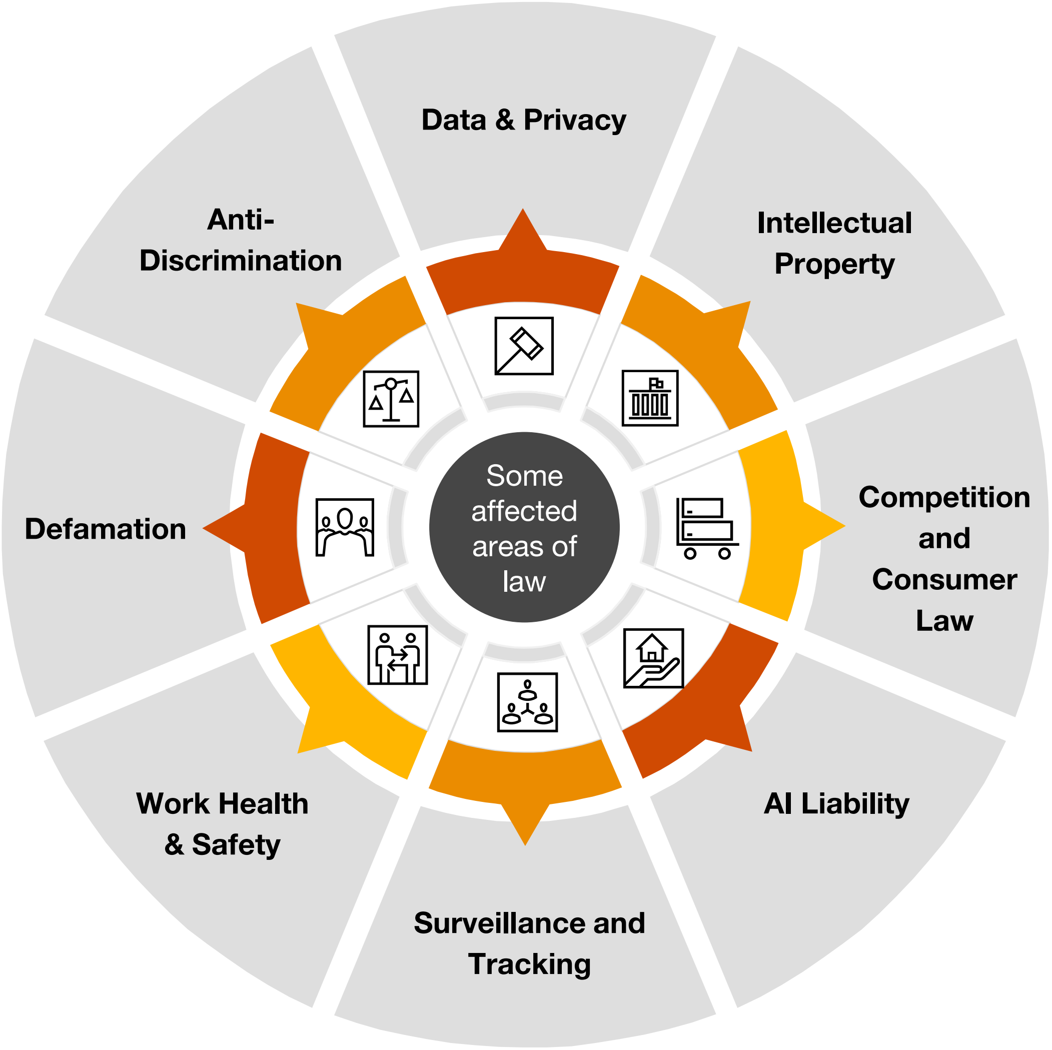
As AI continues to evolve at an extraordinary pace, it is more important now than ever before to recognise the breadth of legal issues under existing Australian legal regimes that intersect with, and apply to, AI.

The opportunities of AI come with an array of challenges for both the law and the wider community to consider. We have provided below a panoramic, non-exhaustive overview of existing areas of law that may apply to the development and use of AI solutions.

Any organisation considering the development and use of AI will need to carefully consider how these regimes will apply.

Governments looking at regulating AI in a meaningful way need to look at existing regimes and how they may apply to AI. Some of these laws may need amending and some may simply need the government or regulators to provide clear guidance on the application of those laws to AI.

Without clear, sensible and appropriate regulation by parliament, there is uncertainty around the legalities of certain AI development and use. Further, in a common law system, (like Australia), without this legislation, these complex new and unexplored implications would be left to the courts to consider in the context of existing law. This opens up the risk of courts adopting a new interpretation that is inconsistent with safe and responsible use and development of AI, or conversely, over-restricting innovation in the name of safety.



2.1 Data and privacy

There is a slew of privacy and cyber security issues raised by the development, use and implementation of AI. Very little has been said about which organisation in the chain of development of AI should remedy such issues when they arise, opening up an accountability and liability gap. The AI lifecycle is complex and involves many different operators, all of which have the potential to impact the quality of the technology.

This article does not undertake an exhaustive exploration of all privacy and cyber security-related issues in AI; but rather, it is designed to provide an overview and act as a launchpad for further discussion.

Data ownership at law

There are innate issues with the concept of data ownership in itself. There is no uniform definition or framework for data ownership at law, and different jurisdictions may adopt different approaches depending on the nature, source, and use of the data.

Ascribing data ownership rights to data used or generated by AI poses a number of legal challenges, including:

1. Data that is processed, generated or used by AI often involve overlapping interests of different parties. It is also not a clear linear division of rights but rather, it is dependent on the nature of the data, the contribution of parties that provided, created, or controlled the data, and the contractual or legal arrangements between those parties. For example, who owns the data or content that is generated by an AI system that has scraped data from a public source? Who owns the data or content that is generated using non-public data e.g. a subscription-based platform? This will undoubtedly then turn to licensing considerations as well as rights in privacy and confidentiality. As with most AI issues, there is no settled position at law on this point yet
2. Where data used or created by AI is, in fact, capable of being the subject of ownership rights, the question then turns to how those rights can be enforced. Broadly, contractual data and confidentiality and equitable confidentiality obligations and intellectual property rights provide the basis and framework for establishing a data owner's rights to restrict usage and disclosure of data (other than personal data). But even then, how can these data owners control the access, disclosure, transfer, deletion, or modification of data that is contained within a 'black box'?

Privacy considerations in implementing AI

In Australia, personal information is subject to data protection obligations under the *Privacy Act 1988* (Cth) (**Privacy Act**) as well as various state and territory privacy laws, regardless of whether that information is publicly accessible.

There is an immediate privacy issue to deal with as AI systems are built upon data. The more data being accessed and used, the greater the risk of privacy issues arising.

Where an AI developer scrapes content, such as blogs, reviews, conversations, comments, or even social media posts, in the creation of their AI system, they must be careful to not fall foul of the Privacy Act, including:

- Inadvertently collecting Australians' sensitive information without consent
- Collecting personal information by unfair means (such as collecting personal information covertly without the knowledge of the individual)
- Not taking reasonable steps to implement practices, procedures and systems to ensure compliance with the Australian Privacy Principles

When it comes to AI applications like biometric facial recognition tools or smart traffic signals, it will be almost entirely impossible to meet the requisite consent requirements under the Privacy Act prior to collection of data.

While traditional ideas of privacy may be challenged by AI, privacy should not have to give way to AI itself. It is incumbent on governments to allow the harnessing of AI in a way that it prioritises and enables privacy. The Australian Government has signposted that it intends to make mandatory a requirement for privacy policies to include meaningful information on the types of personal information used as part of automated decision making, and have also stipulated certain transparency requirements in relation to how automated decisions are made.

Use of AI by SOCI entities and privacy

Under the *Security of Critical Infrastructure Act 2018* (Cth), the storage, transmission or processing of sensitive operational information outside Australia poses a material risk to cybersecurity and privacy. As such, those entities that are subject to the SOCI Act and the OCI Risk Management Program Rules who choose to implement AI will need to consider how the SOCI requirements apply to any potential offshoring of data used as part of an AI system as well as the wider risks and hazards posed by AI usage.

2.2 Intellectual property

Broadly speaking, there is no form of legal protection for creative works made by AI under current law. Generative AI is only going to get bigger, better, and more powerful. With the emergence of increasingly sophisticated AI technologies, the law is at a crossroads when it comes to the authorship and use of AI-generated copyright and patentable works. Regulatory and legislative bodies around the world are faced with the challenge of how to best update intellectual property laws in a way that promotes the use of this new technology without harming creative ecosystems and economies. Given that one of the key purposes of intellectual property laws is to reward creators for intellectual effort, governments have to walk a fine line in dealing with AI.

The intellectual property considerations surrounding AI can be divided into three key areas.

01. The algorithm or system itself

The first is primarily concerned with the AI algorithm or system itself. This is more easily dealt in the area of copyright as the relevant Act recognises computer programs as protected 'literary work' and it is likely that a court would consider that the AI model is a form of computer program.

Under the *Patents Act 1990* (Cth), an invention is patentable if it is not explicitly excluded (which algorithms are not) by the Act and it:

- Is in a manner of manufacture within the meaning of s 6 of the statute of monopolies
- Is novel and involves an innovative step when compared to the prior art base
- Is useful
- Was not secretly used in the patent area before

For AI systems, there should be no issue with showing its utility and that it was not secretly used in the area. It should also be simple enough to show that it is a manner of manufacture, which in modern law is concerned with the invention being an artificially created state of affairs and of utility in a field of economic endeavour. The satisfaction of the second requirement surrounding novelty and innovative steps may be slightly more difficult, depending on whether the system is simply processing data or whether it involves technological innovation.

There is a potential issue, however, when algorithms end up training and building themselves. Who should be able to file a patent for the improved system? If a person's intellectual property is used to train AI, does this mean that person also has rights to the AI output? As will be discussed later, AI is not recognised as an inventor yet in Australia, so would it potentially be the creator of the base system, or can no patent be filed at all?

02. AI inputs

The main intellectual property issues in relation to inputs for AI are:

1. Whether a developer of an AI system is permitted to use certain material for training purposes.
2. Whether a user of an AI system is permitted to input certain material into the AI system to generate an output, where those materials themselves may be protected by intellectual property laws. This is especially where the system is a generative AI system.

If a piece of work is protected by copyright, the owner of the copyright is automatically granted a set of exclusive rights and copyright is infringed when someone else exercises these rights over the works. The rights include:

- The right to copy/reproduce the work in a material form
- The right to publish the work
- The right to communicate the work to the public.

This raises several questions in the AI context, depending on whether the copyrighted works end up being used to train an AI system and/or used to create the output of the algorithm. When using the works to train an AI system, the question is whether there has in fact been an infringement of copyright. At this stage, there is an argument that an AI model is not actually reproducing, publishing or communicating the work when being trained and therefore not infringing copyright.

A prominent example of this issue is the lawsuit that was been filed by various artists earlier this year against Stability AI, Midjourney, and DeviantArt, for infringing the exclusive rights in their copyrighted works when they used these works to train their AI image generator.[4] Similarly, the work of artist Hollie Mengert was used to train an AI model such that where the correct prompt was used, the output generated would be in her artistic style.[5] Would this still be breaching her copyright, since no one specific work is reproduced in a material form, but rather her style is used? Further, even if it did fall in this category, could any of the fair dealing exceptions (such as research or study) in Australia be used to allow this? What if an AI system was trained on outputs of other AI, would there be a breach of intellectual property rights here if technically AI cannot be an 'author'?

The courts in Australia have yet to come to a clear decision as to whether a generative AI system may use inputs that are protected by copyright law, but users of these systems should be aware that they very well might be infringing on another's intellectual property rights in doing so.

AI & Copyright: as seen in Courts

1. Comic book, Zarya of the Dawn, which features illustrations that were based on text prompts fed to Midjourney, an AI image generator, has been granted limited copyright protection by the US Copyright Office. The author, Kashtanova, was found to be the author of the comic's text as well as the selection, coordination, and arrangement of the written and visual elements, however the images themselves do not garner any copyright protection on the basis they are 'not the product of human authorship'. The Copyright Office also stated that Kashtanova did not exercise sufficient creative control over the output of Midjourney and that there was too much 'distance' between the user's input and the AI's output.
2. In August 2023, the District Court for the District of Columbia upheld the US Copyright Office's denial of copyright registration application for AI generated artwork.[6]It stated that 'human authorship is essential part of valid copyright claim' and that the applicant, Thaler, 'played no role in using the AI to generate the work'. The judge found that the absence of a 'guiding human hand' in the AI-generated artwork's creation disqualified it from copyright protection. The Copyright Office further noted that other AI-assisted material could qualify if a human 'selected or arranged' it in a 'sufficiently creative way that the resulting work constitutes an original work of authorship'. Thaler has indicated a desire to appeal the ruling, so watch this space!
3. In contrast to the US judgment, the Beijing Internet Court endorsed copyright for AI-generated artwork in a landmark ruling in November 2023. However, it is crucial to note that in the Thaler matter, Thaler was trying to recognise the AI itself as the author and not the person using the AI as a tool as author.

03. AI outputs

The final, and perhaps more contentious area of legal concern is that of AI 'data outputs' i.e. material created by non-human AI systems. Legal protection, by way of copyright, automatically subsists over any original creative material. The definition of 'original' is tested when it comes to all things AI and machine learning, which draws on existing material to generate output. In order for a work to be sufficiently 'original' to attract protection under the Copyright Act, there needs to be an exertion of human skill, creativity and 'sweat of the brow'. Clearly, an AI system is not able to exert human skill.

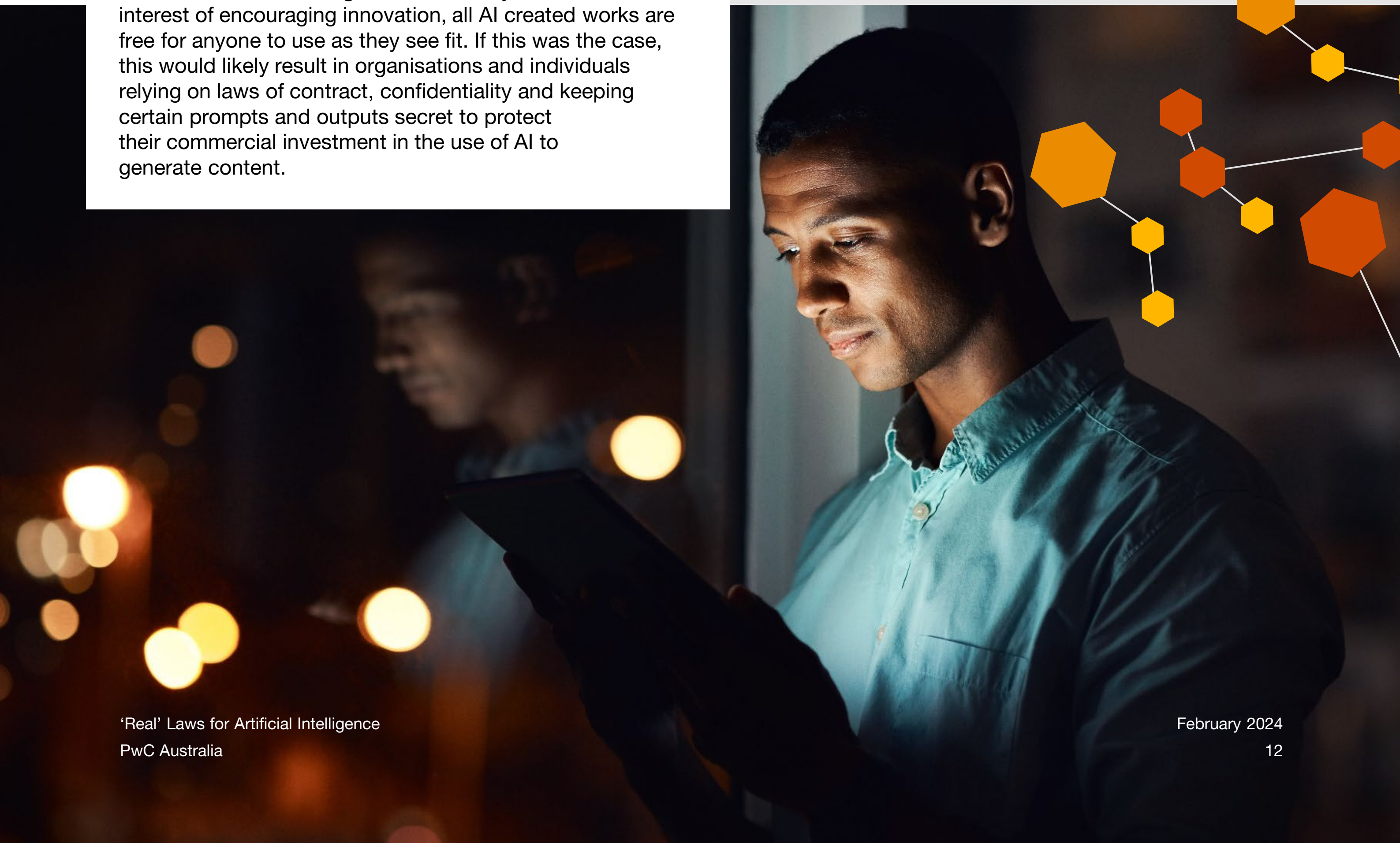
As such, as the law currently stands, an AI cannot itself be credited as the author of any output that it generates. The question then becomes, who is the author? Should the user of the AI system be the author who entered prompts to generate the output? What about the developer of the AI system? What about the original author of the material on which the output is based?

Currently, there is much uncertainty in this space, but a possible approach may involve considering if a person using an AI system is able to show enough creative work in the input/instructions it gives to the AI, that they should be able to claim copyright. A key challenge of this is that as soon as someone enters those prompts to generate an output that is copyrightable, reproduction by someone else of that material with the same prompts will arguably have infringed copyright. Further, if a person is able to claim copyright by virtue of the prompts, then other considerations apply: is it infringement if someone uses the same prompt as another? Will they face responsibility for potential wrongdoing of the AI system if the prompts lead to the copying of a substantial part of copyright works (e.g. copyright infringement)?

For now, this analysis is on a case-by-case basis, but as AI develops and becomes more 'human-like' in its thinking and creations, it may be a very real possibility that AI is able to own its own works – or governments may decree that in the interest of encouraging innovation, all AI created works are free for anyone to use as they see fit. If this was the case, this would likely result in organisations and individuals relying on laws of contract, confidentiality and keeping certain prompts and outputs secret to protect their commercial investment in the use of AI to generate content.

Similarly, in patent law, Australian courts are not yet willing to accept an AI as an 'inventor'. However, it is important to note that Australian courts have drawn a distinction between the question of inventorship and whether an AI-created invention could be patentable. They have said that it is not the fact that these AI-created inventions are not capable of being patented, but rather it is a question of who is able to patent it. Based on the extent to which a person is involved in an invention, they may be able to patent an AI-generated invention on a case-by-case basis. The courts have in fact explicitly given some examples of where a patent may be possible, including if they are the person who inputted the data that the AI used, they developed the AI software, or if they owned the copyright of a source code or the machine running the AI software. Patentability considers two main thresholds.

1. Was an 'inventive step' involved when compared to the existing art base.
2. Was the invention 'obvious' to someone with common general knowledge of that specific technology field. This then takes the question into the realm of whether AI has consciousness to determine whether something was 'obvious' or not. But currently the patentability of an AI generated invention will also be on a case-by-case basis.



2.3
Competition and consumer law

Protecting consumers

Like any product introduced into the Australian market, the Australian Consumer Law (**ACL**) will apply to impose obligations on the supplier of AI products. There are a range of areas where the ACL may impact the development and use of AI.

One such instance concerns the marketing of AI applications. The ACL provides that a person must not engage in misleading or deceptive conduct or conduct that is likely to be misleading or deceptive. It could be very difficult to make accurate representations about AI given the lack of understanding surrounding it and the difficulty in predicting what 'behaviour' an AI system will engage in. Organisations must take care as to how it chooses to use or engage with AI in its operations or there may be significant repercussions. For example, the ACCC recently were successful in their case against the travel booking website, Trivago, for misleading consumers in its use of algorithmic decision making that gave a false impression of providing the best or cheapest deals.[7] Trivago was ordered to pay almost $45 million in penalties for this contravention.

In addition, AI itself has become embroiled in controversy in providing misleading or completely false information when it cannot generate an accurate answer to a query – known as 'hallucinations'. ChatGPT in particular is starting to become infamous for this behaviour. This raises the question: who will be to blame where misinformation is disseminated to consumers?

Similarly, the 'black box' that is AI may make it difficult to establish a breach of a consumer guarantees surrounding safety and quality of products, and identifying an adequate remedy – can you simply repair or replace an AI system with another? Who along the AI lifecycle is responsible and liable for the breach?

AI developers and suppliers will also need to consider how AI will work in light of the **unfair terms regime**, which aims to counter the inherent imbalance in power between consumer and supplier. Such unfair contract terms have recently been made illegal in November 2023 and the penalties have been substantially increased. For individuals, this could mean a penalty up to $2.5m.

Consumers typically do not have much bargaining power in negotiating the terms on which they are sold various products and often have to accept them as presented by the supplier to be able to complete the purchase. Suppliers of AI will need to take care in ensuring that contracts are not too one-sided that they are in breach of the unfair contract term regime, despite how tempting it may be to protect themselves.

Protecting competition

With AI able to ingest and analyse copious amounts of real-time industry data, it has the capacity to affect the competitiveness of market. Without true control or regulation, AI systems can unintentionally (or even intentionally) be used to engage in acts of collusion, especially when market dynamics point towards collusive outcomes being more stable or rewarding. If multiple businesses are using similar AI algorithms, the AI could inadvertently have the effect of price fixing. Beyond collusion, AI could also facilitate exploitation of market power (through discrimination and bias). AI, with no barriers or regulations to prevent it from doing so, could choose to implement and drive anticompetitive strategies, like predatory pricing aimed to drive competitors out of the market, leading to consumer harm. Price discrimination between consumers on a large scale is also foreseeable, with systems being able to use a person's history of spending to predict the maximum price that a particular consumer may be willing to pay, if dynamic pricing is allowed. Travel retailers and sellers on platforms, such as Amazon, are already using algorithms to vary their prices.

There are also prohibitions in the current competition law against abuse of market power. Large businesses with a significant share of the market may be in breach of this if they use large volumes of personal data to train their AI, especially where only they have access to that data due to their market share. Regulators may view this as an unacceptable barrier to entry. Businesses should further be wary of entering into exclusive arrangements with providers of AI if it is not actually necessary for their business model, as they further risk breaching competition law.

This all gives rise to the question of whether AI considerations needs to be addressed in competition law or guidance om the ACCC. Should there perhaps be a prohibition on use of AI that significantly interferes with or lessens consumer rights and market competition?

2.4 Surveillance and tracking

AI is increasingly being used to surveil people in real time, whether it is by law enforcement or by private parties all around the world. Recently, the French government has approved temporary laws that will allow the police to use CCTV algorithms during the Paris 2024 Olympics, allowing them to detect and flag various anomalies, such as crowd rushes, fights or unattended bags.[8] Meanwhile in the UK, a range of spy agencies are lobbying the government to relax surveillance laws (that were enacted following the 2016 leaks from Edward Snowden surrounding state-based surveillance) that they view as a burden on their ability to train AI models with bulk amounts of personal data.

In Australia, surveillance devices are regulated by a regime of various federal, state and territory based legislation, in addition to the Privacy Act. Currently, they cover (at varying levels across Australia) computer tracking/monitoring, as well as optical, audio and workplace surveillance. Critically, these laws do not address facial or other types of facial recognition as of yet.

The Australian Government has been looking to reform this area of law into a modernised legislative framework. It is a very topical issue, with large retailers, such as Woolworths and Bunnings, using AI software '*Auror*' to detect crime and alert security guards in real time to catch shoplifters.[9] Until recently, the Australian Federal Police (**AFP**) was also using Auror to help catch criminal 'gangs' that steal over time at different stores. The AFP recently suspended this use, after a Freedom of Information request revealed that over 100 staff members had been using the platform without consideration of privacy and security implications. Bunnings and Kmart are also currently being investigated by the Office of the Australian Information Commissioner for their use of another facial recognition software application.[10]

It is clear that surveillance and tracking technologies are a significnt source of regulatory concern. Companies must be careful and ensure that any AI based tracking and surveillance they undertake is compliant with laws.

2.5 AI liability

As foreshadowed in section 2.3 of this article above, a key question that needs to be addressed at law is who is ultimately responsible for AI and damage or harm that it causes? The law on this is unclear without clear direction from lawmakers, leaving parties to rely on existing laws in areas such as negligence, consumer law, contracts and corporations law.

There are a number of parties involved in the development and use of an AI system: the data provider, designer, manufacturer, programmer, developer, user, and AI system itself. Each case of failure may require a unique analysis of the surrounding algorithms and circumstances to determine who is at fault. Given this, governments will need to consider whether a streamlined approach to liability may be required. Ultimately, this will depend on which area of law a claim is brought under. For example, in cases where a breach of contract has occurred, the consequences will likely be decided by the allocation of risk within the contract.

This question of liability and accountability will be raised more and more as the propagation of AI increases the harm and damage caused. Set out below are some key areas of law where liability for the operation of AI may come into play.

The tort of negligence

The tort of negligence may assist in determining the liability of AI gone wrong. An example frequently brought up is the case of the self-driving car being tested for Uber that struck and killed a woman.

There were different factors that contributed in the incident, from the algorithm not identifying her quickly enough to recognise the need to break, to the safety operator not paying attention. This incident resulted in the driver being charged with criminal negligence, but had no legal repercussions on Uber as the prosecution office declined to prosecute despite the main cause being a system failure.[11]

The development of common law through cases like that will determine what will constitute a breach of duty of care by the use of AI. Most likely, liability will depend on when someone has a duty of care in using and developing AI, and the foreseeability of it being used the way it was. It is important to note that there are established categories of relationships, such as doctors and patients, where there will automatically be a duty of care already, so these parties will need to take extra care if they decide to make use of AI.

As we track further back the development chain of AI, there will be more issues with proximity and foreseeability. Contributory negligence and vicarious liability may help balance liability between the various parties involved. Damages that may be claimed also has the

possibility of including psychological injury and advice given, which is an important consideration when using AI.

Directors' duties and personal liability

Although specific black letter law regulating AI has yet to be formalised, companies and directors need to understand their role and responsibilities in the deployment of AI.

Under Australian law, a company is a separate legal person to the people running it, protecting directors from personal liability. This corporate veil may, however, be lifted by courts where directors have breached their duties. In that vein, directors could be exposing the company to legal liability if they fail to uphold their statutory duties e.g. acting with reasonable care and diligence, and mitigating preventable harms arising from AI systems created and used by the companies they oversee.

As such, directors and officers must consider how to manage the data, models and people involved in implementing AI, including considering whether AI governance framework/s should be put into place. Appropriate AI governance can, if done correctly, accelerate the growth of a company's uptake and ability to benefit from AI solutions, and ensure directors and officers meet their obligations under the Corporations Act.

See our article, 'AI: What Directors Need to Know' for a detailed breakdown of the relevance of directors' duties to AI and how directors can effectively manage these duties.

Product Liability

AI may be used by retailers and other sellers of goods either as part of the sale process or as the product being sold itself. AI algorithms may contain errors from human programmers which can lead to unintended to consequences. Where there is a claim brought by a consumer under a consumer law related action, consumer liability law should generally decide when a manufacturer or distributer is liable.

One complication arises from the 'black box' nature of AI which make it difficult to fully understand the cause of any malfunction or negative outcome. This leads to uncertainty as to who holds is responsibility. Another difficulty in determining the responsible party is the evolution of AI algorithms, especially where one algorithm in particular is improved or modified along the line, especially by a third party. The ACL is not currently equipped to handle these situations in a way that will result in an equitable, or even sensible outcome. So far, these laws have been useful where automated-decision making (ADM) AI applications have been used, to police any misleading and deceptive conduct.

As discussed later in more detail in section 3.2 of this article, the EU is ahead in this area. The proposed AI Liability Directive introduces a presumption of causality and an information access right for victims. It is also further proposed that the no-fault liability regime in the EU Product Liability Directive be modernised to include AI.

2.6 Employment law

There are a number of ways in which AI may play a role in an employment context. AI has already been used by companies to assist with hiring. AI technologies have been implemented to select, or narrow down, candidates based on historical data of who is the best match for the role, which may lead to discriminatory outcomes.

Employers may also use AI to monitor and direct the work performances of their employees. While there may be an appeal to using AI to track patterns in employee's behaviours, capabilities and habits, it does bring the risk of negatively impacting organisations by introducing greater psychological, structural and physical risks to the people. It is a line that will have to be carefully managed, as it may draw attention to human rights issues and various labour laws, especially around the probity of management decisions. Specialised health and safety training around AI may be necessary for employers to meet their obligations to ensure, as far as reasonably practicable, the health and safety of workers and other people.

In addition, employers may also have to consider the possibility of their employees using AI in a manner aimed at levelling the playing field, especially with their bargaining power. Workers may be able to use AI to observe patterns around hiring and pay e.g. as an extension of already existing platforms such as Glassdoor (especially with certain state laws making salaries non-confidential). Job search websites already often display the data around the salary range of a certain positions at certain seniority levels. Trade unions may also make use of AI to increase their knowledge and power when it comes to collective bargaining and other industrial instruments.

Employees are slowly starting to become empowered to know their worth, how to use it to negotiate better outcomes for themselves, and not accept less – particularly as employment practices become more and more transparent. AI will only enhance this capability, and should be carefully considered in its application.

2.7 Anti-discrimination

Bias in AI systems poses a very real risk of leading to potential breaches of anti-discrimination laws where organisations use these systems to make decisions in relation to individuals. It arises from AI decision-making tools that generate unfair and discriminatory outcomes, often resulting from some form of statistical bias in the underlying training data.

AI is being used more and more frequently by organisations to make important decisions, including by the government. These AI systems are generally trained to apply a probability model based approach to prompts. This often involves utilising historical data to determine the treatment of new data presented to it. The issue arises when the data used to train AI is itself biased or discriminatory, and that is the pattern that the AI will use for future decisions. For example, if an AI is asked to decide on the expected financial success of someone based on being trained on data where there was a gender pay gap, then it will likely conclude that the appropriate amount to pay a man should be higher than a woman.

This can similarly be applied in a criminal context where many countries are already using predictive tools. For more than 20 years, the UK has been using the Offender Assessment System (Oasys) in their criminal justice system to make predictions on matters such as granting bail, the kind of sentence imposed, prison security classification, assignment to rehabilitation and even outcomes of immigration cases. In all this time, no scientist has been permitted access to the data used to independently analyse its working or accuracy. A recent study analysed the predictive performance of such criminal risk assessment tools, and found that the predictive performance was mixed, ranging from poor to moderate. It further found that most validation studies had a high risk of bias, partly due to inappropriate analytical approach being used. One such case where this was apparent was Jordan Sweeney, a convicted murderer that was released from prison in 2022 after being classified as 'medium risk', only to brutally assault and kill a young woman walking home alone after a couple days. A review found that he should have been classified as 'high risk'.

Australia has a large range of anti-discrimination laws, at both the federal and state level, that prohibits discrimination based on age, sex, race and disability – to just name a few of the protected catagories. As outlined in section 2.6 of this article, employers are also becoming more and more reliant on AI technologies and screening tools in the hiring process to pick out the 'best applicant'. These employers need to ensure that they remain abreast of anti-discrimination laws. Even where there is no malice or bad intentions from the employer, the issue is that AI users cannot fully understand the 'black box', and therefore cannot ensure that it is not being used in a discriminatory manner. As such, those looking to use AI systems in this manner must take great care to actively monitor the outputs for indications of unfairness or bias.

Importantly, it appears that AI in the employment context has caught the eye of regulators and governments. The Australian government recently released a consultation paper: '*Updating the Fair Work Act 2009 to provide stronger protections for workers against discrimination'* that looks to integrate more robust protections against bias and discrimination in the workplace. In July 2023, New York City began enforcement of the first-of-its-kind Automated Employment Decision Tool law. The law requires companies who use AI technologies in their hiring and promotion decisions to ensure candidates are aware that an automated system is being used. It also requires these companies to perform an annual audit of their recruitment technology for bias.

2.8 Defamation

Can we rely on anything AI tells us?

The rise of AI has exposed an underlying issue with the algorithms that underpin large language models. That is, does AI actually use accurate and real data in generating a true and useful output, or does it perhaps makes things up – known as 'hallucinating'?

What is defamation?

The action of defamation is concerned around finding a balance between free speech and protecting the reputation of the subject of speech. In Australia, there is an action where there has been communication of a 'defamatory meaning, of and concerning the plaintiff, to a person other than the plaintiff' upon the publication of a defamatory imputation. No proof of actual loss or injury is needed to show damage. A communication will be defamatory if it tends, in the minds of ordinary reasonable people, to injure a person's reputation by:

1. Disparaging them
2. Causing others to shun or avoid them
3. Subjecting them to hatred, ridicule or contempt

The risk brought upon by AI

The main defamation risk in AI arises when AI 'hallucinates' i.e. the AI system makes up its own facts and information with no basis in reality. It is very easy for AI to make false statements or allegations about people in this regard.

The question is: who is responsible for AI generated material published. Liability in ordinary context would generally extend to all involved in the publication (e.g. proprietors, printers and distributors).

There is also a separate question for corporations as corporations are generally exempted from being able to bring a defamation claim. What will they be able to do about false statements made by AI about them that ruins the brand or reputation, impacting on their revenue?

Careful consideration must be had when incorporating AI into defamation laws…

Examples

There are increasingly concerning examples of generative AI hallucinating on serious questions posed by users. Most famously, ChatGPT falsely included a prominent law professor from Georgetown University in a list of scholars that have sexually harassed someone, citing an article from *The Washington Post* that never existed.[12]

In Australia, the mayor of Hepburn Shire Council in Victoria was falsely accused of being a guilty party in a foreign bribery scandal, when in reality he was the whistle-blower who informed the authorities of the bribe payments.[13] He is currently suing OpenAI for defamation – this is an area to watch, as it will likely be a landmark decision.

03

Global AI regulatory approaches

3.0
AI around the world

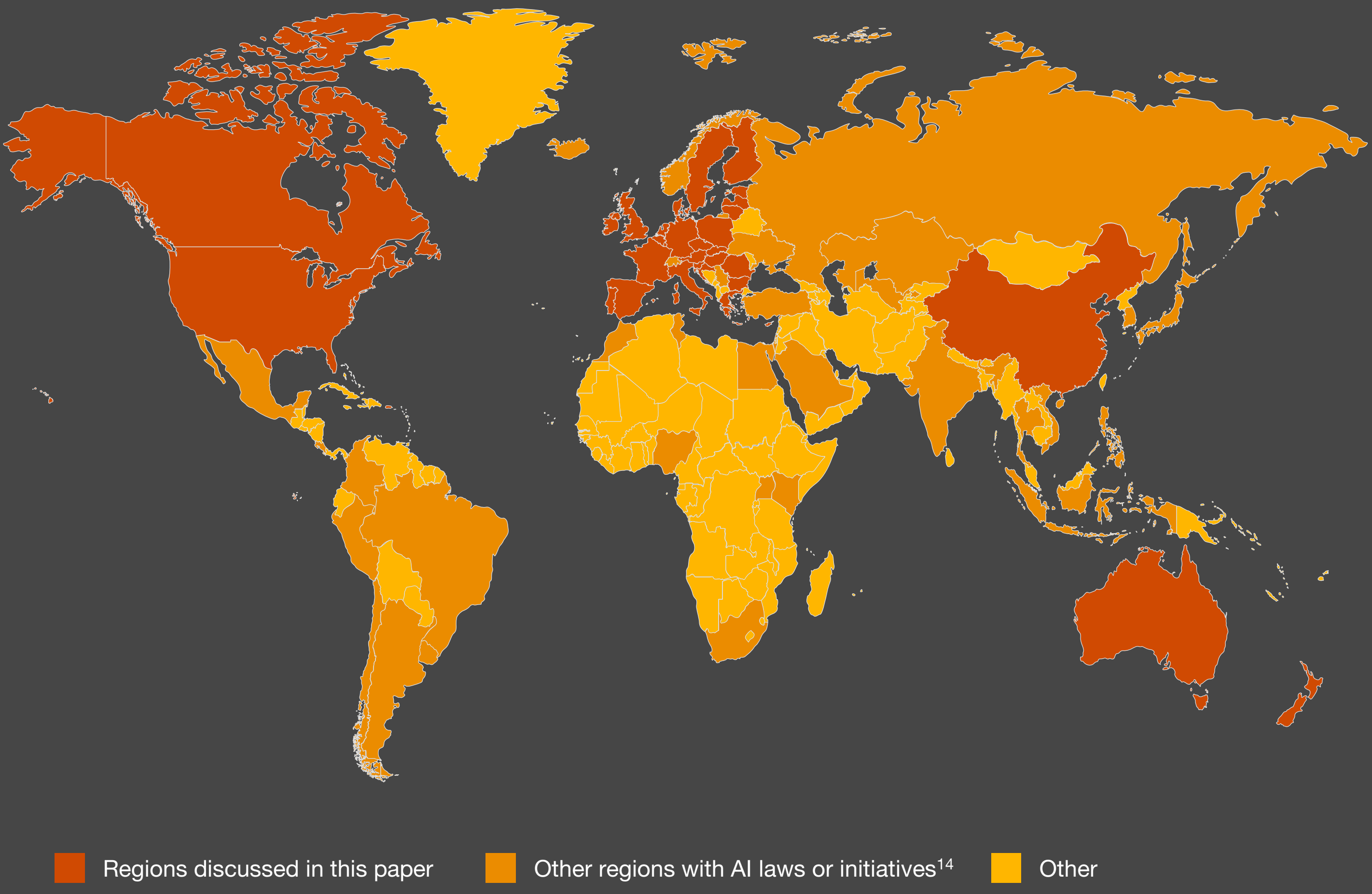
Government across the globe are moving to introduce specific regulations for AI

Given the increased use and capability of AI, governments have recognised an urgent need to regulate AI in order to manage the potential risk of harm that it poses to people and society. The challenge for governments is to find a regulatory approach that still encourages innovation and AI development. In truth, an appropriately structured regime which strikes this balance could in fact accelerate the responsible growth of AI within a jurisdiction.

Currently there is no globally agreed approach on how AI should be regulated. In the east, China has moved quickly to establish rules for Generative AI (see section 3.5 below). In the west, Europe has perhaps led the pack in proposing a risk based regulatory framework but even its legislative process has not been able to accelerate the AI Act into law (we now expect that this may not be in place until 2025). Some countries have already established draft legislation on AI, while others have only provided cursory wide-sweeping statements around AI best practices. Some countries have yet to even address it publicly, opting for a 'wait and see' approach.

In this section, we have summarised the current state of specific AI use and development regulation in Australia and select jurisdictions around the world. This is by no means an exhaustive list of the jurisdictions that are moving to regulate AI.

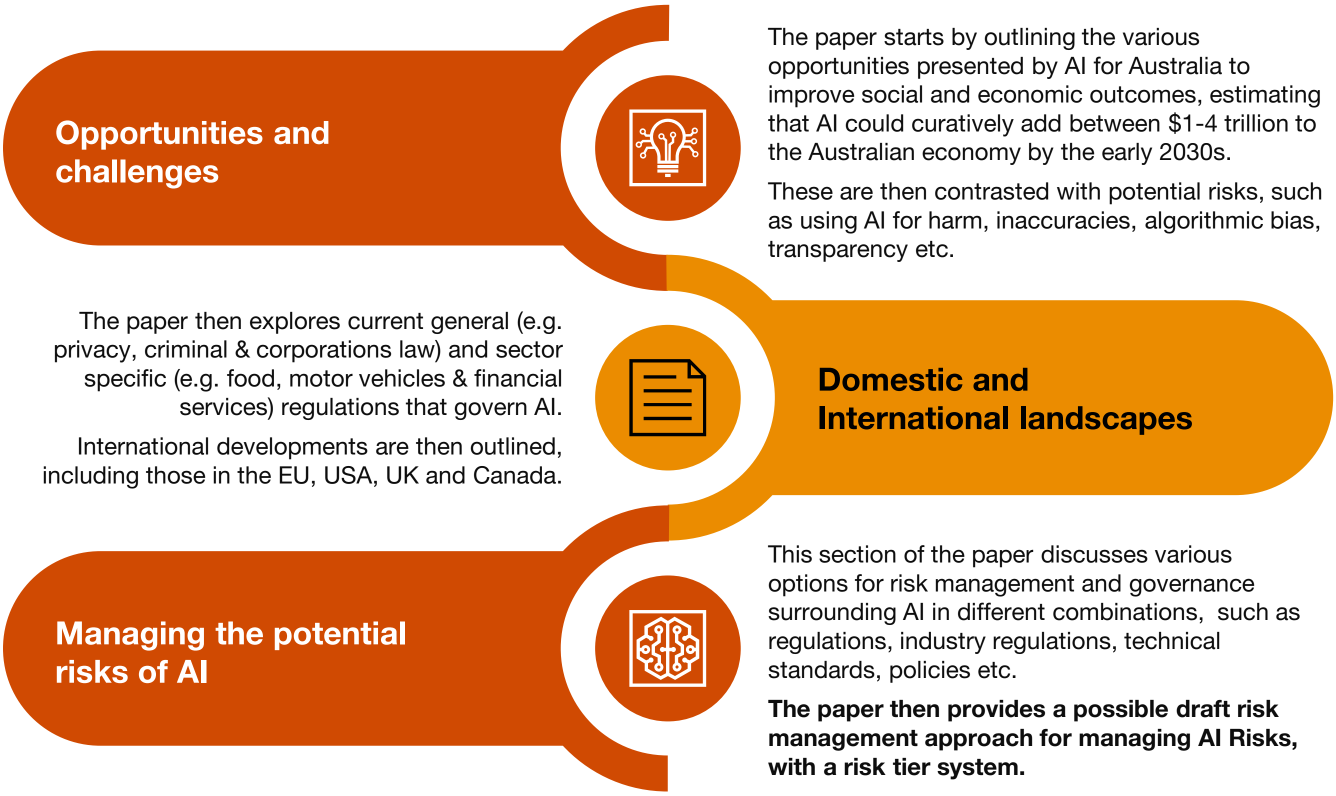
The world is moving to regulate AI…


🟥 Regions discussed in this paper
🟧 Other regions with AI laws or initiatives[14]
🟨 Other

3.1
Australia

To date, Australia has no overarching or specific legislation that deals with either the use of or market surrounding AI. The Minister for Industry and Science announced in early 2024 that legislation will be fast tracked to regulate AI in high risk settings. A new advisory body to government will work with industry, academics and government to establish this legislative framework and define the types of 'high risk' technologies and applications that will be captured be under the law.

In June 2023, the Government also released the ‘Safe and Responsible AI in Australia’ discussion paper, which provides an overview of AI opportunities and risks, existing domestic governance, Australia’s broader regulatory framework, recent and ongoing international developments and invites feedback on a number of questions concerning whether further governance and regulatory responses are needed in Australia. These are more specifically outlined below. The deadline for this consultation was in early August 2023. Our submission to the consultation is linked at section 5 of this article.



Opportunities and challenges

The paper starts by outlining the various opportunities presented by AI for Australia to improve social and economic outcomes, estimating that AI could curatively add between $1-4 trillion to the Australian economy by the early 2030s.

These are then contrasted with potential risks, such as using AI for harm, inaccuracies, algorithmic bias, transparency etc.

Domestic and International landscapes

The paper then explores current general (e.g. privacy, criminal & corporations law) and sector specific (e.g. food, motor vehicles & financial services) regulations that govern AI.

International developments are then outlined, including those in the EU, USA, UK and Canada.

Managing the potential risks of AI

This section of the paper discusses various options for risk management and governance surrounding AI in different combinations, such as regulations, industry regulations, technical standards, policies etc.

The paper then provides a possible draft risk management approach for managing AI Risks, with a risk tier system.

The Federal Government released its interim response to the Safe and Responsible AI in Australia consultation on 17 January 2024. It is clear from both the Government response and the industry submissions that a key priority is to ensure that regulatory measures adopted encourage the safe development and deployment of AI systems, but do not interfere with low risk innovation in Australia’s domestic tech sector (including trading and export activities) and the ability to take advantage of AI-enabled systems supplied on a global scale. The Government will look to work with industry partners and continue to engage internationally to build appropriate protections and ‘guardrails’ to manage AI risks throughout the AI lifecycle, from the design stage to development and eventual deployment. The consultation made clear that **both regulatory and non-regulatory initiatives** were necessary to mitigate AI risks (both emerging and existing).

Regulatory action

Hybrid approach in regulating

It appears that Australia will likely take a two-pronged approach to regulating safe and responsible AI:

1. Updating/adapting over 10 existing legal frameworks to regulate some risks of AI e.g. privacy, IP, anti-discrimination, competition and consumer laws; and
2. Establishing ex-ante regulation (i.e. specific AI regulation), particularly for the deployment of AI systems in legitimate, but high-risk, settings and for 'frontier' or 'general purpose' AI models. These regulations will prescriptively deal with monitoring, design and deployment of AI.

Risk-based approach

The Government's initial response commits to a 'risk-based' approach that is capable of responding to AI concerns even as the landscape continues to shift, similar to the EU AI Act.

'High-risk' AI systems

Taking another leaf out of the EU's book, Australia is also looking to introduce a concept of 'high-risk' AI with a 'systemic, irreversible or perpetual' impact – these AIs will be subject to a more regimented set of rules and compliance obligations.

Technology neutral regulation

Australia will consider a 'technology-neutral' or 'outcomes focused' approach to regulating, in effort to work around the fast-paced advancements in AI and avoid unnecessary or disproportionate burdens on businesses and regulators.

Security

Security was identified as a necessary foundation to build community and business trust in AI. Amongst other measures, Australia will honour the Bletchley Declaration and support global action to ensure AI models are secure by design.

Non-Regulatory Action

AI advisory body

Many submissions called for the establishment of an AI advisory body, which the Government has heeded temporarily. An advisory committee, consisting of experts, will be established to guide the development of mandatory guardrails for high-risk AI.

Regulatory sandboxes

Introduction of an Australian AI regulatory sandbox could allow the government to work closely with industry developers to test and trial new AI concepts/technologies in a monitored environment – providing the ability to grow domestic AI capability. The idea of a 'sandbox' has been accepted by EU legislators and we will soon see it take shape in the EU AI Act.

Continued investment in AI

The 2023-24 Budget already contains over $75 million of funding for pure AI initiatives. The Government will continue to consider opportunities to support the adoption/ development of AI and other automation technologies.

Building trust in AI

To generate trust in AI, the Government will consider opportunities for safe and responsible adoption and use of AI technologies e.g. the development of practical guidance and educational initiatives to help the public understand AI better, take a 'community-first' view, and welcome public involvement and technical expertise in developing AI guidance/laws.

International initiatives

Australia is committed to aligning with and supporting international partners in shaping global AI regulation and governance.

While consultations will continue to establish mandatory safeguards, especially for high-risk applications of AI and watermarking of AI-generated ADM. The Government has indicated its immediate priorities are (1) developing a voluntary risk-based AI Safety Standard, (2) developing options for voluntary labelling materials, and (3) establishing an expert advisory body to support the development of options for further AI guardrails. Australia will continue to engage internationally to help shape global AI governance, and also monitor how its international counterparts are responding to the challenges of AI to ensure our domestic responses are interoperable with global processes.

There are also a range of AI-specific plans and frameworks that provide guidance to how organisations may consider governing their use of AI. The most important of these may be the 2019 Australian AI Ethics Framework, which provides 8 principles to guide businesses and governments navigating AI.

Principles Guiding AI – Australian AI Ethics Framework	
1. Human, Societal and Environmental Wellbeing	AI systems should benefit individuals, society and the environment.
2. Human-Centred Values	AI systems should respect human rights, diversity, and the autonomy of individuals.
3. Fairness	AI systems should be inclusive and accessible and should not involve or result in unfair discrimination against individuals, communities, or groups.
4. Privacy Protection and Security	AI systems should respect and uphold privacy rights and data protection and ensure the security of data.
5. Reliability and Safety	AI systems should reliably operate in accordance with their intended purpose.
6. Transparency and Explainability	There should be transparency and responsible disclosure so people can understand when they are being significantly impacted by AI and can find out when an AI system is engaging with them.
7. Contestability	When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.
8. Accountability	People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

Similarly, New South Wales released an AI Assurance Framework in March 2022 which, along with their mandatory AI Ethics Policy, forms a major component of their AI Strategy. It sets out a mandatory review process, consisting of a self-assessment as per the prescribed questions and submission to the AI Review Body, that all departments and agencies within New South Wales must use when developing, building and implementing AI projects. The Framework focuses on the five mandatory principles for the use of AI.

Community Benefit

AI should deliver the best outcome for the citizen, and key insights into decision-making.

Fairness

Use of AI will include safeguards to manage data bias or data quality risks.

Privacy and Security

AI will include the highest levels of assurance.

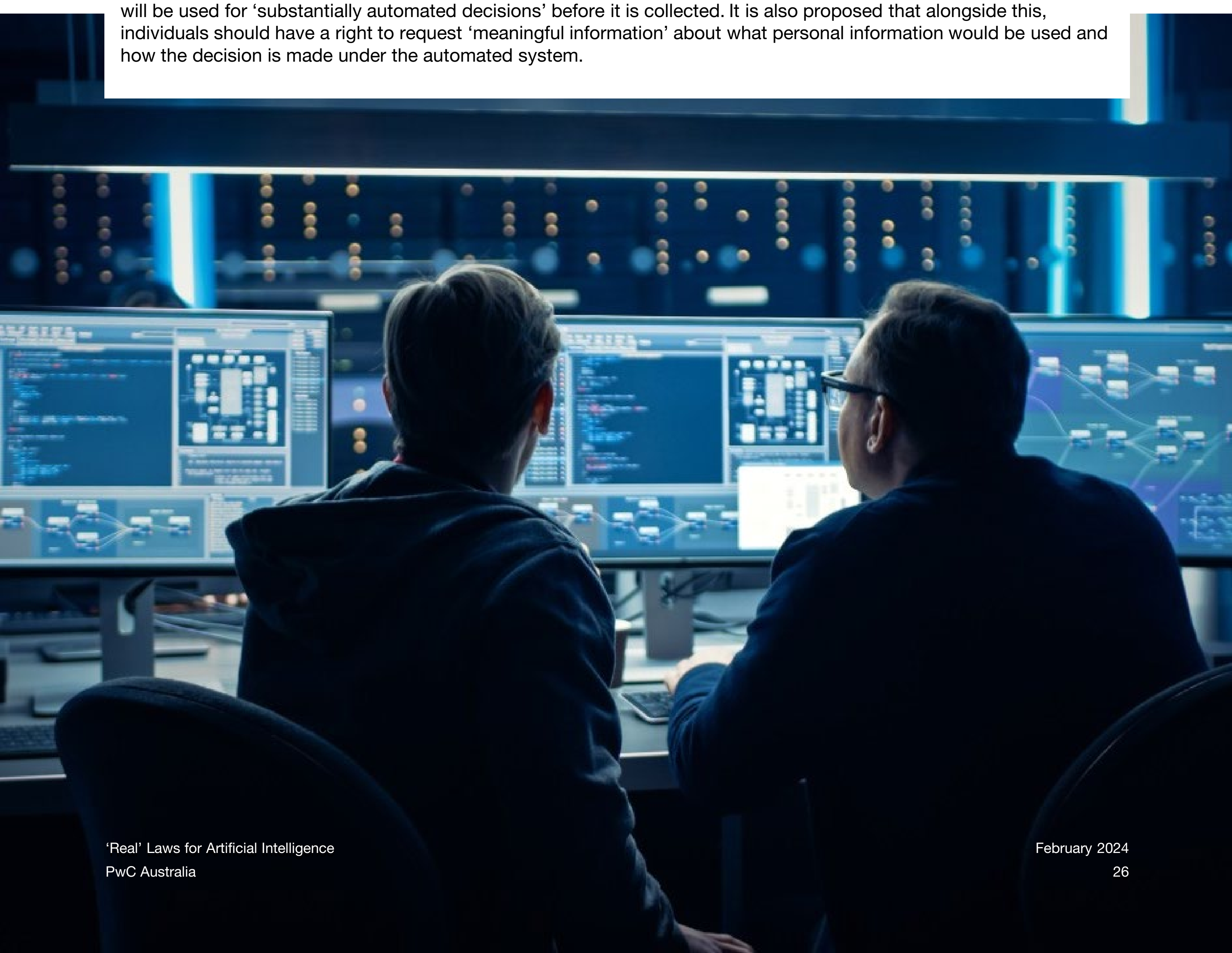
Transparency

Review mechanisms will ensure individuals can question and challenge AI-based outcomes.

Accountability

Decision-making remains the responsibility of organisations and individuals.

It is yet to be seen if other states will follow this example and which direction Australia will take. Users of AI should also consider the existing legislation that may extend to managing AI related risks in some circumstances, such as *the Privacy Act 1988* (Cth), *Online Safety Act* 2021, and the Australian Consumer Law. There is already an indication that the Privacy Act will be reformed to incorporate greater individual protections regarding automated decision making. A recent review into the Privacy Act recommended that entities should notify individuals that their personal information will be used for ‘substantially automated decisions’ before it is collected. It is also proposed that alongside this, individuals should have a right to request ‘meaningful information’ about what personal information would be used and how the decision is made under the automated system.



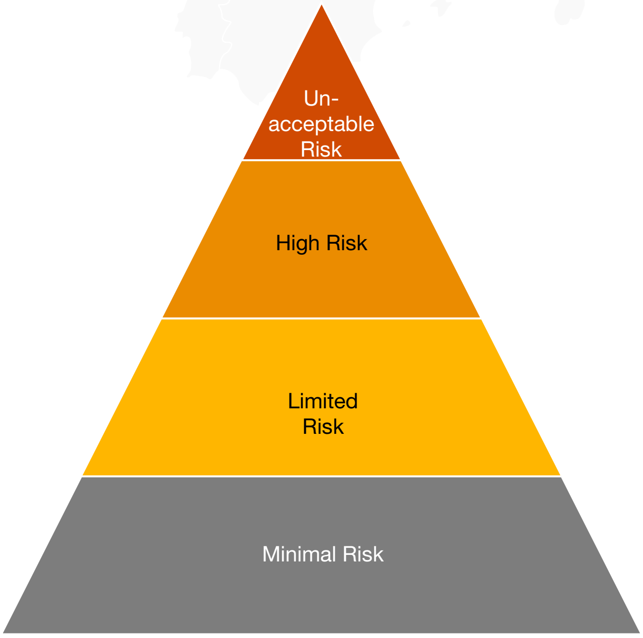
3.2
The European Union

Currently, the EU has some regulations surrounding the use of automated decision making in the General Data Protection Regulation (GDPR). For example:

- Individuals cannot be subject to certain fully automated decision making, which must be safeguarded against by data controllers.
- Where automated decision making does operate, the individual must be informed of this when their data is collected and how this will affect them.

The EU is, however, in the process of implementing a targeted dual approach in regulating AI systems. **There is firstly the AI Act**, focused on controlling the AI systems that can come onto the European Market and under what conditions. In the past couple of years, both the European Parliament and the Council of the EU have each adopted their proposals and recently undertook trilogue negotiations to agree on a position. After months of debate and a three day prolonged meeting to finalise the position, an agreement was reached on 8 December 2023. The Council of the EU released the text of the provisional agreement on 2 February 2024. With approval confirmed from the Council of the EU, we will now look to the European Parliament as they begin their approval processes. Significant opposition is not expected.

The second is the **introduction of the AI Liability Directive** and **amending the Product Liability Directive**, to target who is responsible for malfunctioning AI and compensating for the harms caused.



AI Act

The AI Act approach provides a classification framework, where systems are labelled one of four risk categories, which will dictate the level of restrictions it faces to be put on the EU market.

Generally, the lower risk category includes systems that the draft Act does not impose obligations on, such as spam filters and video games. There are then specific systems that pose a ‘limited risk’ that face certain transparency requirements, where users must be informed that they are interacting with AI and what the AI is doing. Much more onerous obligations are imposed on ‘high risk’ systems, that can have a significant impact on the life of a user. A large list of examples is provided, including real-time and ‘post’ remote biometric identification systems without agreement, systems for critical infrastructure and law enforcement, and systems influencing employment and migration, to name a few. If a company fails to comply with the following obligations, they could face fines up to the higher of €35,000,000 or 7% of the company’s worldwide turnover for the preceding financial year:

- Requirements around high-quality data, testing for robustness and accuracy, transparency, adequate human oversight, and appropriate documentation practices;
- Conformity assessments to pass legal obligations, bearing the CE logo and registration on the EU database, initially and anytime significant changes are made to the system; and
- Establishment of AI risk management processes.

There is also an additional category, for systems that pose ‘unacceptable risk’ and are completely prohibited from the EU Market. Examples include biometric categorisation systems that use sensitive characteristics, untargeted scraping of facial images from the internet or CCTV footage to create facial recognition databases, emotion recognition in the workplace or educational institutions, systems used for social scoring and systems that exploit vulnerabilities of specific groups or deploying subliminal techniques, or that manipulate human behaviour to circumvent free will.

A series of safeguards and narrow exceptions have recently been negotiated in the new December deal around the use of biometric identification systems (**RBI**) in public, for a strict list of crimes with judicial authorisation. ‘Post-remote’ use will only be able to be used to conduct a targeted search of a person convicted or suspected of having committed a serious crime. In real time, RBIs only can be used for limited locations for a limited period of time, with struct conditions, only for the purposes of targeted searches of certain victims, prevention of specific and present terrorist threats and the localisation or identification of a person suspected of having committed one of the specific crimes mentioned in the regulation.

The deal has also brought about a separate set of obligations for all general purpose AI (GPAI), with its own classification framework which will identify some AI as ‘high impact with systemic risk’. Some of the obligations for GPAI include maintaining technical documentation to help downstream providers comply, complying with EU copyright law and publishing a statement about the data used to train the algorithm. For GPAI that meet the criteria for high-impact with systematic risk, they will have to conduct model evaluations, assess and mitigate systemic risks, conduct adversarial testing, report to the Commission on serious incidents, ensure cybersecurity and report on their energy efficiency. After the harmonised EU standards are published, such GPAIs will be able to rely on codes of practice to comply with the regulation. Unlike other AI systems that will be enforced nationally, GPAIs will be enforced by the Commission’s AI Office. To encourage innovation in the field of AI, the EU AI Act also allows for ‘regulatory sandboxes and real world testing’ before AI products are placed on market.

Liability Directive

The AI Liability Directive has two primary features. The first is the presumption of causality, which will shift the burden of proof for victims of AI systems where:

- The conduct of the developer or deployer of the AI fails to meet a duty of are directly intended to protect against the harm that occurred under a EU or national law.
- The failure influenced the functioning of the AI.
- The output, or failure thereof, of the AU was reasonably likely to have caused the damage.

The satisfaction of these three limbs reverses the onus and provides a rebuttable presumption that the AI caused the harm. Where the AI system is classified as ‘High Risk’ under the AI Act, there is a list of actions that automatically trigger a breach of duty of care under the first limb. The second feature is an access right, allowing victims to request disclosure of information about high-risk AI, to help with the usual difficulty associated with gathering evidence in civil law systems.

It is also proposed that the no-fault liability regime in the EU Product Liability Directive be modernised, to include AI systems, AI enabled goods and software in the definition of ‘products’, to account for modern digital characteristics in deciding defectiveness, allowing compensation for loss of data and introducing new rules, allowing for compulsory discovery for liability claims.

AI Pact

To help bridge the transitional period before the AI Act is enforceable, the AI Pact was announced as an initiative encouraging ‘the voluntary commitment of industry to anticipate the AI Act and to start implementing its requirements ahead of the legal deadline’. Companies and other organisations will take a pledge to work towards compliance with the upcoming AI Act and details of what actions they will be taking to achieve this. The Commission will collect and publish these, and then gather the interested parties in the first half of 2024 to discuss how to best further the pact.

3.3 Canada

In June 2022, the Artificial Intelligence and Data Act (AIDA) was tabled by the Canadian government, as part of Bill C-27, the Digital Charter Implementation Act. The AIDA outlines a framework for a new regulatory model, which will be built on and evolve through an 'open and transparent regulatory process', with further consultation with stakeholders. It intends on being a risk-based approach that aligns with the EU AI Act, the OECD AI Principles, and the US NIST Risk Management Framework. The current intention is for the AIDA to fill in the gaps of the current regulatory landscape. With the various development stages discussed, it will not be in force until at least 2025.

There are three main areas the AIDA focuses on. The first of these is building on existing Canadian human rights and consumer law by ensuring that 'high-impact systems' are held to the same standards to 'which Canadians are accustomed'. Canada's department for regulating industry and commencer, the Innovation, Science and Economic Development (ISED) has provided guidance on how AIDA could look at regulating these systems –it has suggested that what qualifies as 'high impact' will involve a consideration of various factors such as severity of potential harms, scale of use, imbalances of economic or social circumstances, or age of impacted persons, harm it has already caused, if people can opt out, if risks are regulated under other laws, and risks to health and safety, or adverse impact on human rights.[9]

Some examples have also been provided, including biometric systems used for identification and inference, systems critical to health and safety, systems that can influence human behaviour at scale and screening systems impacting access to services or employment.

If a system is classified as high-impact, appropriate measures would need to be implemented to identify, assess and mitigate risks of harm or biased output, before the system can be made available for use. The obligations imposed by the AIDA will focus on human oversight and monitoring, transparency, fairness and equity, safety, accountability and validity and robustness.

The second aim of the Act is to empower the relevant Minister to administer and enforce the Act through a new office headed by an AI and Data Commissioner, to ensure the policy and enforcement evolves with the technology.

The AIDA will be looking to create new criminal provisions that would prohibit reckless and malicious uses of AI that could cause serious harm. There are also severe monetary penalties proposed for violations of the AIDA, as well as the Administrative Monetary Penalties (AMPs) for violation of the an Act, regulation or by-law that still need to be drafted.

3.4 The United States

The United States does not have a comprehensive federal legislative approach to regulating the use of AI, however the White House recently issued an Executive Order (discussed in detail on the next page), which directs various US authorities to begin developing an approach to development, deployment and use of AI.

The US has also established a few voluntary frameworks for the use and development of AI. The latest is the AI Risk Management Framework (RMF) released by the National Institute of Standards and Technology. The goal of the AI RMF is to offer a resource to the organisations that are designing, developing, deploying, or using AI systems to help integrate risk management principles into the lifecycle of AI systems. The RMF is intended to be voluntary, rights-preserving, non-sector-specific, and use-case agnostic, providing flexibility to organisations of all sizes and in all sectors and throughout society to implement the approaches in the RMF. This is discussed in more detail in section 4.1 of this article.

Broadly, it appears that AI regulation in the US will be a state-based issue. This individualistic state-by-state approach is similar to how privacy regulation has been developed within the US (due to the lack of overarching federal legislation in the subject matter).

However, with that being said, there has been some movement on the national front. The National Telecommunications and Information Administration (NTIA) has invited public response to four key areas to support its AI-related work and provide a starting point to the approach that should be taken by the US. The Federal Trade Commission (FTC) is also looking to do some work here. In January 2024, the FTC made it very clear that model-as-a-service companies (i.e. companies that develop and host AI models to make available to third parties via an end-user interface or an API) that fail to abide by their privacy commitments to their users and customers, may be liable under the laws enforced by the FTC.

Most recently, the Biden-Harris administration secured ‘voluntary commitments’ from seven leading AI companies regarding the safe, secure and transparent use of AI technology.

As indicated above, AI regulation in the US has come via state-specific action (as opposed to a nation-wide initiative). Various states having a range of proposed, pending, and enacted AI legislation, usually focusing on discrete issues.

For example, many states have legislated (or are attempting to legislate) on the use of AI in the employment context. A Californian bill proposed to prohibit the use of automated-decision-making systems during the hiring process if the systems discriminated based on protected characteristics. New York City began enforcement in July 2023 of its Automated Employment Decision Tool law, which requires companies who use AI and other machine learning technology as part of their hiring process to notify candidates that an automated system is being used. The law also mandates that an annual audit of their recruitment technology must be performed, with a view to check for bias.

Illinois has successfully enacted the Artificial Intelligence Video Interview Act, which requires employers who solely rely on these systems to further a candidate in the hiring process to report its use to the government data, allowing them to determine whether the AI discloses a racial bias in its use.

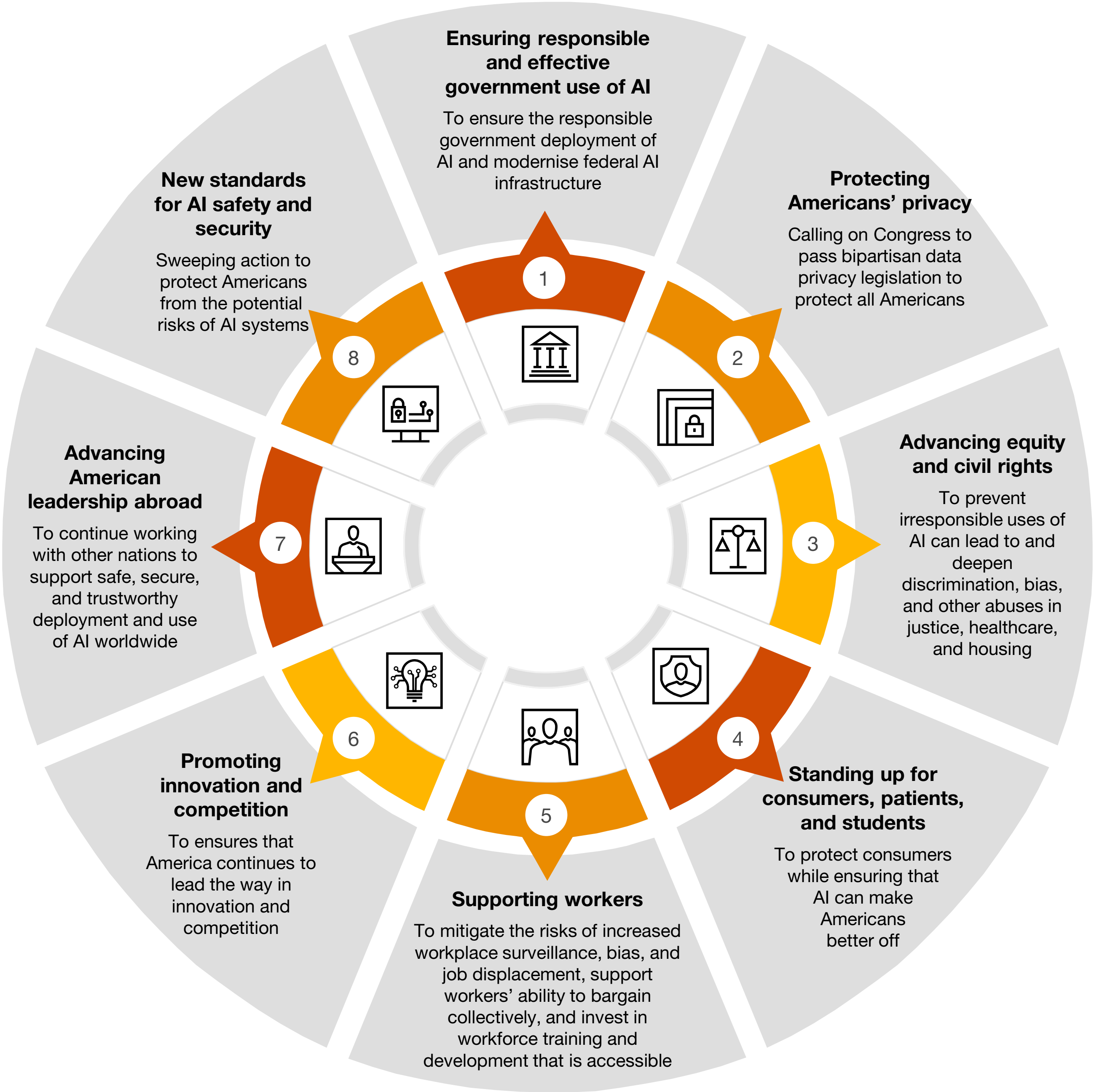
Many states are also focusing on introducing a range of commissions and committees to support and investigate AI in different contexts, particularly the use of AI in government decision-making. California is looking to introduce their own ‘Office of Artificial Intelligence’ for this purpose. Massachusetts, New York and Rhode Island are following California’s example in establishing commissions in this area. These are just a few examples of how individual states are attempting to integrate AI regulation into existing laws.

In any event, the US has recognised the need to position itself as a leader in trustworthy, inclusive, and responsible AI. In doing so, the National AI Advisory Committee (NAIAC) was formed. It first convened in May 2022, and consists of 26 leading AI experts in private sector, academia, non-profits, and civil society, with the purpose of providing the President guidance in this area.

Executive order

Most recently, the President issued an a new executive order on 30th October 2023 on Safe, Secure and Trustworthy Development and Use of Artificial Intelligence. It directs Federal US Agencies to adopt practices to ensure that they are using any AI in a responsible way, that will ensure the protection of Americans. The Executive Order may have a wider implications, such as flow down of obligations from these agencies to private parties, like developers. In effect, it could become a non-binding minimum standard. It may also have a significant impact on the development of AI regulation worldwide.

The order directs the following actions:



Ensuring responsible and effective government use of AI

To ensure the responsible government deployment of AI and modernise federal AI infrastructure

Protecting Americans' privacy

Calling on Congress to pass bipartisan data privacy legislation to protect all Americans

Advancing equity and civil rights

To prevent irresponsible uses of AI can lead to and deepen discrimination, bias, and other abuses in justice, healthcare, and housing

Standing up for consumers, patients, and students

To protect consumers while ensuring that AI can make Americans better off

Supporting workers

To mitigate the risks of increased workplace surveillance, bias, and job displacement, support workers' ability to bargain collectively, and invest in workforce training and development that is accessible

Promoting innovation and competition

To ensures that America continues to lead the way in innovation and competition

Advancing American leadership abroad

To continue working with other nations to support safe, secure, and trustworthy deployment and use of AI worldwide

New standards for AI safety and security

Sweeping action to protect Americans from the potential risks of AI systems

3.5 China

From early on, China has reiterated its strategy to promote the healthy development and technology innovation in the area of AI. In July 2017, China released a '*New Generation AI Development Plan*', and has since published a number of regulatory instruments on the path of implementing this Plan. The approach tends towards targeting specific algorithms and systems specifically, rather than the area of AI as a whole.

Most recently, on 13 July 2023, the Cyberspace Administration of China (**CAC**), published new 'interim' rules targeting generative AI, following the draft version from April 2023. The initial version specifically covered 'models and related technologies' used to generate content, suggesting that companies execute contracts with users and safeguard the content generated. There was also a focus in the draft to ensure the generative content uphold the 'core value of socialism', 'respect social morality and public order', and does not attempt to 'subvert state power' or 'undermine national unity' or produce content that is pornographic, or encourages violence, extremism, terrorism or discrimination. The focus of the final rules included:

1. Respect intellectual property and fair competition.
2. Protect personal information (e.g., obtain data subject consent or apply other legal basis).
3. Formulate clear and operable labeling rules.
4. Protect user's input record and do not collect excessive info.
5. Provide complaint channel. These rules came into effect on 15 August 2023.

Some of the regulatory instruments preceding this include

Regulations on the Administration of Deep Synthesis of Internet-Based Information Services (January 2023)

Opinions on Strengthening the Ethical Governance of Science and Technology (March 2022)

Provisions on the Management of Algorithmic Recommendations in Internet Information Services (March 2022)

Ethical Norms for New Generation AI (September 2021)

China has also recently unveiled its 'Global AI Governance Initiative' (GAIGI) which aims to shape AI regulation for participating countries, and has the potential for significant implications on how different countries approach AI regulation, especially when it comes to issues such as export controls and global AI supply chains.

3.6
Other regulatory approaches

There are several jurisdictions that have considered AI and come to the decision that regulating it through legislation is not the correct approach, at least not yet.

New Zealand

New Zealand is currently piloting a policy project titled '*Reimagining Regulation for the Age of AI*' alongside the World Economic Forum as part developing a national AI Strategy. It is aimed at co-designing an actionable governance for AI regulation in the future by inviting a national conversation about obtaining a social licence for the use of AI, developing an in-house understanding of AI for well-informed policies, and mitigating the risks surrounding AI. However, there has been minimal movement otherwise from the New Zealand government.

The United Kingdom

The UK has been relatively tentative in its approach to regulating general purpose AI thus far. Following the National AI Strategy (proposed in 2021), the UK government released a white paper 'A pro-innovation approach to AI regulation' in March 2023. This provided a framework based on five principles which is proposed to be implemented through a non-statutory basis with existing regulators having the statutory obligation to implement them. The government released its response to the paper on 6 February 2024, confirming its plans to introduce the 5 principles as well as a context-specific framework and voluntary measures for AI developers to consider. It intends to consult on its plan throughout 2024.

The UK also hosted an international AI Safety Summit late 2023, attended by various AI industry, policy and research experts. It made headway in the diplomatic space with a joint commitment by twenty-eight governments – including US, EU China, Brazil, India, and Indonesia – and leading AI companies to put advanced AI models through a series of safety tests before release. The UK also announced the creation of an AI Safety Institute. However, interestingly, it appears that actual AI regulation in the UK may still be a while away as the UK's first minister for AI and intellectual property publicly stated in November 2023 that there would be no UK law on AI 'in the short term' as the government was concerned that heavy-handed regulation could curb industry growth.

Singapore

Singapore has no current plans to introduce regulation specific to AI. Instead, regulators in Singapore have adopted a 'light-touch' approach which focusses on developing a range of guidance and principles that seek to enable the safe use of AI.

Singapore's high-level strategy for AI is outlined in the National AI Strategy (last updated in 2023). It outlines the country's plan to be at the 'forefront of development and deployment of scalable, impactful AI solutions' and raise up individuals, businesses, and communities to use AI with confidence, discernment, and trust. In 2018, Singapore established the Advisory Council on the Ethical Use of AI & Data for the purpose of advising and supporting the government on the ethical use of data-driven technologies.

Singapore has developed a Model AI Governance Framework (last updated in 2020), which is intended to promote public understanding and trust in AI. The framework covers traditional AI models and provides guidance to assist organisations in addressing key ethical and governance issues that exist with AI systems. In recognition of the rapid advancements of generative AI, Singapore has announced at the World Economics Forum in Davos its draft Model AI Governance Framework for Generative AI, which it hopes to finalise in mid-2024. The framework expands on the original Model AI Governance Framework.

A lynchpin in Singapore's AI governance strategy is AI Verify, a governance testing framework and toolkit for organisations that want to conduct a voluntary self-assessment of their AI systems. The system helps companies attempt to objectively and verifiably demonstrate to stakeholders that their AI systems have been implemented in a responsible and trustworthy manner.

04

Getting ahead: Governing AI in your organisation

4.0 Governing AI

Whilst there is risk in adopting AI, often the bigger risk is not adopting AI and falling behind

Organisations should be looking for ways to accelerate the implementation of AI quickly, but responsibly, and stay ahead of the curve to stay competitive and compliant. It's about moving forward with both 'veracity' and 'velocity'.

This includes establishing a clear, holistic and robust governance framework to underpin the development, implementation, procurement and use of AI technologies, in-line with best practice.

There are a growing number of AI governance and risk management frameworks, standards and principles published by both public and private sector entities (including the EU AI Act and the Australian Government's own Discussion Paper that sets out a framework for AI risk management), which can form a useful base for any AI risk management framework.

While there are a number of thematic similarities across the various frameworks in terms of key principles such as ethics, fairness, data and security, the appropriate framework for your business will differ depending on your organisation's scope/use of AI tools and its risk appetite.

This section explores how organisations can get ahead of the curve and implement flexible AI governance processes and procedures that can be easily aligned to any future regulatory regime that may apply.

4.1 Governance frameworks

Establishing an effective AI governance framework should be a priority for companies and their directors as the use of AI systems becomes more and more widespread. An AI governance framework is shaped by an organisation's appetite for risk and therefore overlays risk management frameworks.

AI governance frameworks are crucial in ensuring the appropriate development, implementation, procurement and use of AI technologies by an organisation. A good governance framework sets out the roles and responsibilities relating to the performance and ongoing monitoring of AI systems, performance metrics, internal organisational structures and accountability for AI outcomes.

Guidance on effective AI governance

Research from the Human Technology Institute reveals that good AI governance can be achieved by implementing the following considerations:

1. Company directors being agile by creating and updating an AI strategy.
2. Designing and implementing clear document systems, policies and standards and delegations for a comprehensive governance framework.
3. Greater awareness of the potential risks for different AI systems.

However, empirical research also reveal that developing ethics-based principles alone is not sufficient for an AI governance framework. Unlike regulations, guidance material or enforceable undertakings, ethical principles are often toothless and tend to assert statements to the consumers and wider public that are not reflected in the governance systems.

ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system

Entities that provide or use AI systems should have regard to the newly published ISO/IEC 42001, which sets out requirements for establishing, implementing, maintaining, and continually improving an Artificial Intelligence Management System (AIMS) within organisations. The standard provides guidelines for the deployment of applicable controls to support AI management processes, e.g. determination of company objectives, third party management, management of risks and opportunities, etc.

Three Lines of Defence (3LoD) model and AI

The traditional 3LoD risk management model is a risk management framework that assists organisations in assign and coordinate risk management roles and responsibilities – it is considered best practice in many industries.

Different groups within an organisation are given a distinct role within this model:

L1 – The first line of defence is the function that owns and manages risks. It is generally management staff that have the primary responsibility of identifying and managing risk as part of their day-to-day operational activities. Other accountabilities assumed by the first line include design, operation, and implementation of controls.

L2 – The second line function specialises in compliance/management of the risk itself. This comes in the form of frameworks, policies, tools, and techniques to support risk and compliance management in the first line.

L3 – This function provides objective and independent assurance. The responsibility for third line is to assess whether the first and second line functions are operating effectively. They report to the board and audit committee, in addition to providing assurance to regulators and external auditors.

The question is whether the 3LoD framework, in its current form, is relevant and effective in managing risks that emerge from AI use and development. A proposed approach is set out below

L1 – This function will be provided by the accountable owner of the AI system and they will be both accountable and responsible for assessing, identifying and responding to the AI risks.

L2 – The second line function will advise and guide L1 on their aligned and compliance with required AI policies and obligations.

L3 – This will be an independent assurance team that the entity's AI policies, frameworks and controls are appropriate and effective in design and operation.

There is already a wide range of guidance materials and template frameworks that can assist organisations with beginning their journey with AI governance or enhancing their existing AI governance mechanisms. Here are some key examples:

Existing governance frameworks

1. **AIGA AI Governance Framework**

   The AIGA AI Governance Framework provides a template for directors and other decision-makers to ensure a practice-oriented framework for implementing responsible AI and adopting a systematic approach for AI governance.

2. **Singapore Model AI Governance Framework**

   The second edition of this Singaporean framework usefully provides real-world illustrations of companies implementing sections of the framework. In particular, this framework provides an approach for companies to consider how a desired probability or severity of harm can determine the level of human involvement in AI-augmented decision-making.

3. **Artificial Intelligence Assurance Framework**

   Proposed by the NSW Government, this AI Assurance Framework balances the benefits and risks of using AI. After answering a series of questions on benefits, risks and ethical principles, a risk rating is assigned. This rating will determine when projects should stop or proceed with or without risk mitigations.

4. **Australian Government Department of Industry, Science Energy and Resources – 8 Ethical Principles for Safe and Responsible AI in Australia**

   The Australian Government Department of Industry, Science Energy and Resources has developed 8 voluntary ethics principles designed to build public trust in your company and positively influence outcomes from AI. Implementing these principles into practice promotes fairness, protection, and security within your company is key to a director’s duty in exercising their powers and perform their functions with care and diligence.

4.2
Risk management frameworks

Risk governance is a cornerstone to ensuring AI is developed, implemented and managed appropriately. In fact, it is perhaps only with proactive risk mitigation measures that AI may truly be accepted by the broader community/ As we all are aware, there are many people who view AI with a lens of scepticism and/or fear. If strong risk management regimes are in place, doom-sayers, cynics and alarmists may be assuaged, and trust will be established ultimately enhancing the utility and viability of AI solutions and investment.

Many governments and organisations have begun to build out risk management frameworks for AI. These include, amongst others:

- ISO/IEC 23894:2023 – Information technology – Artificial intelligence – Guidance on risk management (further described below)
- EU AI Act risk based regulation approach
- NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0) (further described below)
- US Department of Energy – AI Risk Management Playbook
- Microsoft Responsible AI Standards
- AIGA AI Governance Framework
- PwC Responsible AI Framework
- Deutschland AI Cloud Service Compliance Criteria Catalogue (AIC4)

Deutschland AI Cloud Service Compliance Criteria Catalogue (AIC4)

BSI's AIC4 provides AI-specific criteria, which enable an evaluation of the security of an AI service across its lifecycle. The criteria set a baseline level of security to ensure that the AI service provider uses proper processes and controls, which can be reliably assessed through independent auditors. It has been specifically developed for AI services that are based on standard machine learning methods and then iteratively improve their performance by utilising training data. The BSI have stated that this criteria make up the minimum requirements for professional AI usage.

PwC has developed and execute assurance procedures against the compliance criteria and have been able to provide reasonable assurance in the form of a ISAE3000 Assurance Report.

In recognition of the burgeoning use of AI by organisations in the market, the International Standards Organisation also published a ISO/IEC 23894 in February 2023 to assist in AI risk management. This Standard provides guidance on how organisations can manage risk specifically related to AI, including the integration of risk management into AI-related activities and functions. The Standard is not mandatory, however it is a possibility that the Australian government, as we have seen with the *Security of Critical Infrastructure Act 2018* (**SOCI**) (Cth), may decide to point to the Standards in the legislation itself at some point, thereby enshrining it in law for certain entities.

ISO/IEC 23894:2023

Information technology – Artificial intelligence–Guidance on risk management

This is a relatively new standard that offers strategic guidance to organisations that 'develop, produce, deploy or use products, systems and services' using AI, for managing the related risks across the lifecycle of the AI. It also offers assistance on the integration of risk management for AI related activities and functions.

The standards are divided into 3 parts:

1. **Principles** – Describes the underlying principles of risk management, including the sources of risk and the considerations that must be taken into account for each of these.
2. **Framework** – Outlines the framework and its purpose, being to assist in integrating risk management into organisation's significant activities and functions, focusing on development, provisioning or offering, or use of AI systems.
3. **Processes** – Provides risk management processes that involve the systematic application of policies, procedures and practices to the activities of communicating and consulting, establishing the context, and assessing, treating, monitoring, reviewing, recording and reporting risk.

NIST Framework

The NIST Framework is a voluntary resource, that organisations can use in the designing, developing, deploying or using of AI systems to 'help manage the many risks of AI and promote trustworthy, responsible development and use of AI systems'.

It is the predominant guiding document that provides practical information/directions to organisations on the capabilities that it should consider to ensure the responsible and trustworthy design of AI and appropriate management of AI risks.

1. Framing Risk

Understanding and Addressing Risks, Impacts and Harms + Challenges for AI Risk Management

- **Risk Measurement**: risks related to third-party software, hardware and data, tracking emergent risks, availability of metrics, risks at different stages of the AI lifecycle, risk in real-world settings, inscrutability and human baseline
- **Risk Tolerance**: the readiness of the AI actor to bear the risk for its objectives
- **Risk Prioritisation**: being realistic about what risks may be eliminated and their urgency
- **Organisational Integration and Management of Risk**: AI risk management should be integrated and incorporated into broader enterprise risk management strategies and processes.

2. Audience

The need for a broad range of perspectives and actors across the AI lifecycle where there will be a diversity of experience, expertise, and backgrounds and comprise demographically and disciplinarily diverse teams.

3. AI Risks and Trustworthiness

The characteristics of trustworthy AI and guidance on how to address them:

- Valid and reliable
- Safe
- Secure and resilient
- Accountable and transparent
- Explainable and interpretable
- Privacy-enhanced
- Fair – with harmful bias managed

4. Effectiveness of the AI RMF

The expected benefits for users of the framework and the need for periodic evaluations of whether the AI RMF is improving the actor's ability to manage AI risks, including policies, processes, practices, implementation plans, indicators, measurements, and expected outcomes.

5. AI RMF Core

The core of the RMF provides outcomes and actions that enable dialogue, understanding, and activities to manage AI risks and responsibility develop trustworthy AI systems. This involves four functions:

- **Govern:** A culture of risk management is cultivated and present
- **Map:** Context is recognised and risk related to the context are identified
- **Measure:** Identified risks are assessed, analysed and tracked
- **Manage:** Risks are prioritised and acted upon based on a projected impact.

6. AI RMF Profiles

Implementations of the AI RMF functions, categories, and subcategories for specific settings or applications based on the requirements, risk tolerance, and resources of the Framework user (e.g. use case profiles and temporal profiles).

4.3 Responsible AI

It is also critical, when discussing the use of AI in organisations, to have conversations and strategise how to design, develop and deploy AI responsibly, in a way that fosters trust and develops confidence. Beyond the previously discussed legal risks of using AI, there are several other categories that should be taken into consideration, so that the organisation is able to respond to queries from various stakeholders that will be concerned. As outlined in the PwC Responsible AI Toolkit, the main risks that should be considered include

Performance

- Risk of errors
- Risk of bias and discrimination
- Risk of opaqueness and lack of interpretability
- Risk of performance instability

Economic

- Risk of job displacement
- Enhancing inequality
- Risk of power concentration within one or a few companies

Security

- Adversarial attacks
- Cyber intrusion and privacy risks
- Open source software risks

Societal

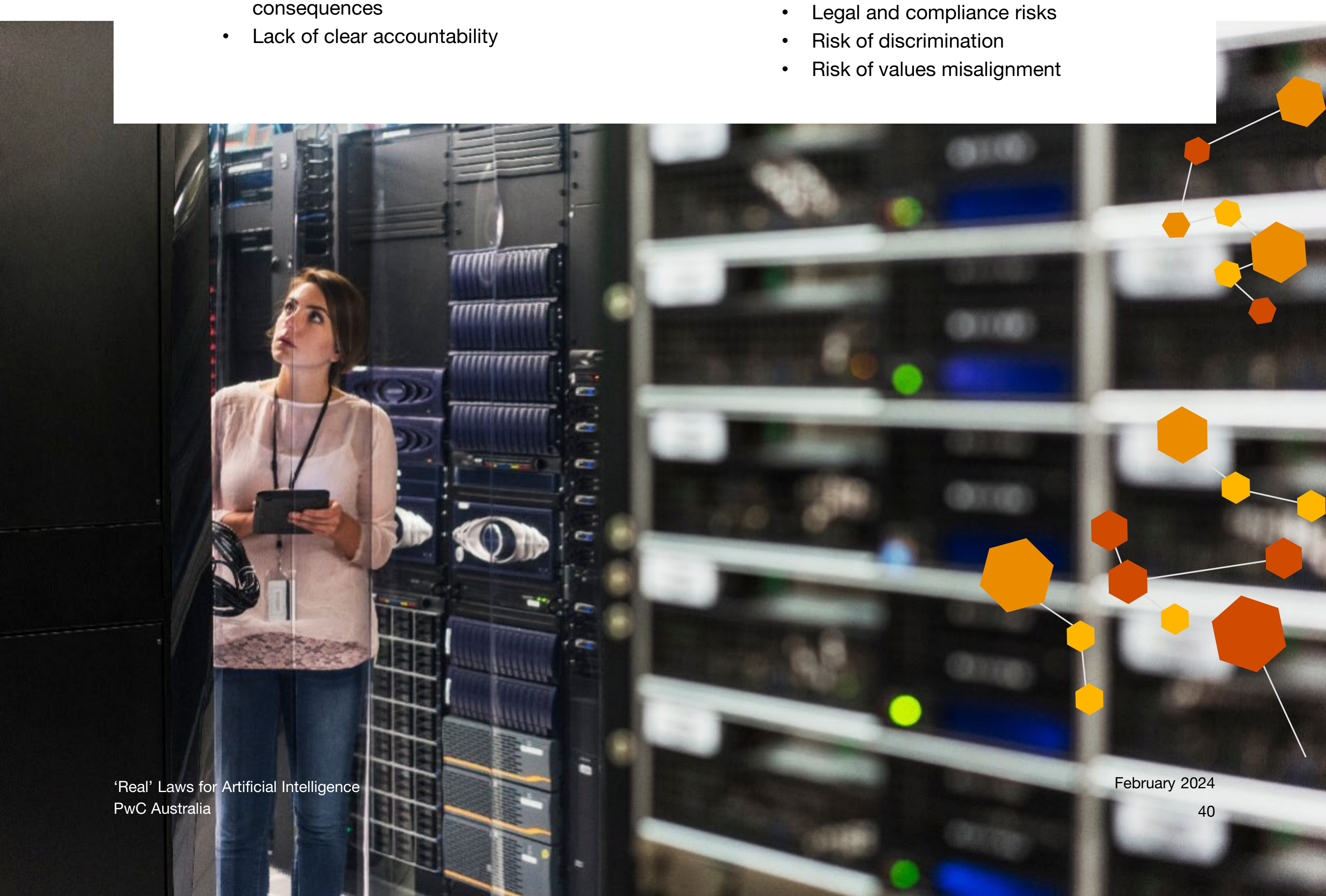
- Risk of misinformation and manipulation
- Risk of intelligence divide
- Risk of surveillance and warfare

Control

- Lack of human agency
- Detecting rogue AI and unintended consequences
- Lack of clear accountability

Enterprise

- Risk to reputation
- Risk to financial performance
- Legal and compliance risks
- Risk of discrimination
- Risk of values misalignment



05

Useful resources

For more information on managing and implementing AI in your organisation, please consider these resources:

- *‘Artificial Intelligence: What Directors Need to Know’*
- PwC Responsible AI Toolkit
- Developing your organisation’s AI policy
- Managing the risks of generative AI
- Unlocking the benefits of AI in the enterprise
- Safe and responsible AI in Australia
- How generative AI can help improve business

Scan to visit our artificial intelligence microsite.

Wherever you are, we'll start there. Together.

No matter where you are in your AI journey, we're here to help, hand in hand. With over 20 years of experience as a leader in technology and data governance, coupled with deep expertise in cloud and digital technologies, and a diverse global ecosystem of alliances, we've assembled the blend of skills and experience that you need to accelerate responsibly.

Mainstream use of artificial intelligence (AI) exploded onto the scene with ChatGPT and given the myriad of commercial applications for generative AI, it is looking like it is very much here to stay. As a result, many agencies and businesses are looking to embed AI into their day-to-day operations. But in amongst the plethora of legal, commercial and risk issues related to AI, where do you start?

PwC's Digital, Cyber and Technology Law team is a team of specialist commercial, technology and intellectual property lawyers. We provide market leading solutions to help our clients solve their most complex information technology and legal problems.

Contact us to discuss how PwC can assist you in preparing your business for your AI journey.

Key contacts

Adrian Chotar
Partner | Head of Digital, Cyber and Technology Law
PwC Australia
T: +61 (0)457 808 068
E: adrian.chotar@au.pwc.com

James Patto
Director | Digital, Cyber and Technology Law
PwC Australia
T: +61 (0)431 275 693
E: james.patto@au.pwc.com

Tom Pagram
Partner | AI Leader
PwC Australia
T: +61 (0)451 470 509
E: tom.pagram@au.pwc.com

David Ma
Director | AI Risk & Governance Lead
PwC Australia
T: +61 (0)413 564 828
E: david.ma@au.pwc.com

Authors and contributors

James Patto
Director | PwC Australia
Digital, Cyber & Technology Law

Annie Zhang
Manager/Associate | PwC Australia
Digital, Cyber & Technology Law

Khushboo Ruhal
Lawyer | PwC Australia
Digital, Cyber & Technology Law

Endnotes

[1] Microsoft, ‘Medicine Man: How AI is bringing humanity back into healthcare’ (Feature, 11 July 2018).

[2] Julie Steenhuysen, ‘Study finds Google system could improve breast cancer detection’ *Reuters* (online, 2 January 2020) <https://www.reuters.com/article/us-health-mammograms-ai-idUSKBN1Z0206>.

[3] Jake Rudnitsky and Mark Bergen, ‘Musk, Wozniak urge halt to more powerful AI models’ *Australian Financial Review* (online, 30 March 2023) <https://www.afr.com/technology/musk-wozniak-call-for-halt-on-developing-more-powerful-ai-models-20230330-p5cwhs>.

[4] Shanti Esccalante-De Mattei, ‘Artists File Class Action Lawsuit Against AI Image Generator Giants’ *ARTnews* (online, 17 January 2023) <https://www.artnews.com/art-news/news/artists-class-action-lawsuit-against-ai-image-generator-midjourney-stability-deviantart-1234653892/>.

[5] https://waxy.org/2022/11/invasive-diffusion-how-one-unwilling-illustrator-found-herself-turned-into-an-ai-model/

[6] *Stephen Thaler v. Shira Perlmutter and The United States Copyright Office* (1:22-cv-01564) (June 2, 2022)

[7] *Australian Competition and Consumer Commission v Trivago N.V.* [2020] FCA 16.

[8] Lisa O’Carroll, ‘French court’s approval of Olympics AI surveillance plan fuels privacy concerns’ *The Guardian* (online, 18 May 2023) <https://www.theguardian.com/world/2023/may/18/french-courts-approval-of-olympics-ai-surveillance-plan-fuels-privacy-concerns>.

[9] James Vyver, ‘Australian retail giants and police using artificial intelligence software Auror to catch repeat shoplifters’ *ABC News* (online, 10 June 2023) <https://www.abc.net.au/news/2023-06-10/retail-stores-using-ai-auror-to-catch-shoplifters/102452744>.

[10] Josh Taylor, ‘Bunnings and Kmart halt use of facial recognition technology in stores as privacy watchdog investigates’ *The Guardian* (online, 25 July 2022) <https://www.theguardian.com/technology/2022/jul/25/bunnings-and-kmart-halt-use-of-facial-recognition-in-stores-as-australian-privacy-watchdog-investigates>.

[11] Aarian Marshall, ‘Why Wasn't Uber Charged in a Fatal Self-Driving Car Crash?’ *Wired* (online, 17 September 2020) <https://www.wired.com/story/why-not-uber-charged-fatal-self-driving-car-crash/>.

[12] Pranshu Verma and Will Oremus, ‘ChatGPT invented a sexual harassment scandal and named a real law prof as the accuse’ *The Washington Post* (online, 5 April 2023) <https://www.washingtonpost.com/technology/2023/04/05/chatgpt-lies/>.

[13] Laura Mayers, Stephen Martin, and Debbie Rybicki ‘Hepburn mayor may sue OpenAI for defamation over false ChatGPT claims’ *ABC News* (online, 6 April 2023) <https://www.abc.net.au/news/2023-04-06/hepburn-mayor-flags-legal-action-over-false-chatgpt-claims/102195610>.

[14] OECD.AI (2021), powered by EC/OECD (2021), database of national AI policies, accessed on 31/01/2024, <https://oecd.ai>.

Thank you

A community of solvers coming together in unexpected ways to solve the world’s important problems

www.pwc.com.au

At PwC Australia, our purpose is to build trust in society and solve important problems. PwC is a network of firms in **152 countries and almost 327,000** people who are committed to delivering quality in assurance, advisory and tax services. Find out more and tell us what matters to you by visiting us at www.pwc.com.au
PWC200871347