Artificial intelligence and your organisation

Legal information for community organisations

This fact sheet covers:

► what is artificial intelligence?

► the benefits of using AI

► legal risks associated with use of AI

► ethical concerns associated with use of AI

► how does AI impact organisational governance?

► implementing responsible AI practices

In this fact sheet, we consider the legal implications of the use of artificial intelligence (AI) on community organisations.

! Disclaimer

This fact sheet provides general information about the legal implications of the use of artificial intelligence on your organisation. This information is a guide only and is not legal advice. If you or your organisation have a specific legal issue, you should seek legal advice before deciding what to do.

Please refer to the full disclaimer that applies to this fact sheet.

There are benefits to community organisations using AI. If used correctly, AI can improve an organisation's work processes and help the organisation meet community needs.

Your community organisation should be aware of the limitations and risks associated with use of AI and take action to manage this.

What is artificial intelligence?

Artificial intelligence or 'AI' refers to technology that can simulate human intelligence, such as learning and problem-solving. AI can refer to different technologies capable of performing complex tasks like learning from experience, adapting to new information, understanding natural language, recognising patterns, and making decisions.

We used to think these tasks were so complex that they required a human brain. However, with the rapid development of technology, computer systems can now perform those tasks in way which can appear human-like.

When most people use the term AI today, they are not referring to the sentient machines from science fiction (sometimes called 'true' AI). Instead, they are referring to computers which simulate or mimic human thought processes.

Depending on what the AI does, there are different types of AI, including machine learning AI, deep learning AI, natural language processing AI, and generative AI.

What is machine learning AI?

Machine learning AI programs 'learn' from data to improve their performance over time. These programs are given access to data, and instructed or programmed to recognise and extract patterns from that data.

Machine learning AI programs are often used to predict consumer behaviour, and suggest or recommend products or services (for example, on social media). They do this by collecting and analysing consumer behaviour like spending habits and location and demographic information, and then grouping customers based on that data.

Examples of machine learning AI include:

- recommendation systems – these systems recommend products, services, or content to users based on their past behaviour or preferences (such as recommendation systems are used by Amazon, Netflix, and Spotify)
- fraud detection – machine learning can be used to identify fraudulent activity, such as credit card fraud or identity theft
- spam filtering – machine learning algorithms can be used to identify and filter spam emails

A not-for-profit could use a machine learning system to analyse donor data to predict future giving patterns, allowing the organisation to target fundraising efforts more effectively.

What is deep learning AI?

Deep learning AI is an advanced type of machine learning AI.

Deep learning AI uses a series of complex algorithms designed to simulate the human brain which means they can analyse larger amounts of data in more complex ways and draw more nuanced predictions and conclusions.

Deep learning AI is used in autonomous or self-driving vehicles, which must analyse a constant stream of data (live footage of the surrounds) and make complex decisions in real time based on that data.

What is natural language processing AI?

Natural language processing AI are programs which can understand, generate and manipulate human language as it is commonly written or spoken ('natural' language).

Examples of natural language processing AI include:

- machine translation systems – these translate text from one language to another
- chatbots – these are computer programs which allow humans to interact with those devices or programs in a seamless, natural way (such as ChatGPT)
- voice assistants – such as Apple's Siri and Amazon's Alexa, can understand and respond to spoken language
- computer vision – this type of AI allows computers to see and understand the world around them and is used in applications such as facial recognition, self-driving cars and medical diagnosis

A not-for-profit could use natural language processing to automate tasks like analysing social media sentiment or translate donation requests into multiple languages.

What is generative AI?

Generative AI are programs that generate new content (images, text, audio, video), typically based on instructions or 'prompts' from a user. They are able to do this by learning patterns from existing data or content and reproducing that content in different forms or combinations.

The most well-known type of generative AI is ChatGPT (the G in ChatGPT stands for 'generative'). By typing out a question or request, users can 'chat' with ChatGPT on a wide range of topics. ChatGPT responds by generating text responses based on the question or request, and the existing data it has access to.

A not-for-profit might use generative AI to produce marketing materials or create educational content.

AI is a rapidly evolving field, and new applications are being developed all the time. You can even use AI to help you find other AI tools (for example, theresanaiforthat.com).

The benefits of using AI

The use of AI has revolutionised various sectors by enhancing efficiency, accuracy, and decision-making processes. These activities range from automating routine tasks to analysing vast amounts of data swiftly and accurately to provide insights that might otherwise remain hidden.

By incorporating AI strategically, not-for-profit organisations can streamline operations, make better use of resources, and increase impact in their communities.

Caution

This next section outlines some of the key benefits and use cases for AI, with reference to specific AI tools. Not all tools are available in Australia, and it is important to bear in mind that **every** AI tool has limitations and risks – including IP infringement, security and data privacy-related risks.

Improving and automating processes

One of the main attractions of AI for organisations of all sizes is the ability to improve and automate processes and operations, freeing up people's time and attention to concentrate on other tasks. This can be even more important for organisations which do not have the resources to engage someone with a particular skill set or capability like marketing or grant writing.

These improvements can range from routine administrative tasks like sorting emails, scheduling meetings, taking notes and processing donations, through to grant and proposal writing tailored to the opportunity.

Similarly, for those organisations with a large volunteer base and multiple events, AI scheduling tools can help to match volunteers with suitable roles based on their availability and skill sets.

Research

There are a range of AI tools which can be used to improve an organisation's research capabilities. These tools can analyse and collate large amounts of data from a range of sources, including websites, academic articles and research papers, social media and more. This research can then be used to generate content, aid decision-making and planning, and streamline operations.

For example, tools like Scite and Consensus can connect and reference over 1 billion different academic citations and can provide an indication as to whether an article supports, mentions or contradicts an assertion or other work.

Similarly, PDFgear Copilot and PDF.ai can analyse large PDF documents (or groups of documents) and present the information in a way that is easily searchable and digestible. Some of these tools also have a ChatGPT-like function where you can ask questions about the content of the PDF and have the information presented to you in a conversational way.

Using these tools can help accelerate the research process and make researching more intuitive and streamlined. However, as with all AI tools, it is important to independently verify all research with reference to the source material such as the original article or paper.

Virtual assistants

Virtual assistants, like chatbots, can greatly enhance the interaction between not-for-profits and their stakeholders by handling routine enquiries and providing general information about the organisation and its activities. They can also monitor and manage online donation platforms and initiatives, ensuring that they are always available for beneficiaries and donors.

One of the significant advantages of AI is the ability to be continuously available. Virtual assistants can operate 24/7 without fatigue, meaning that the organisation is always 'available' to stakeholders and can provide timely responses to any enquiries.

However, virtual assistants will not be suitable for all organisations and uses.

What about using virtual assistants for meetings?

Most workplace software now includes tools (virtual assistants) to help with meeting tasks, like transcribing meetings in real-time or translating a meeting into another language (for example, Microsoft Teams and Zoom).

When evaluating whether a virtual assistant is a good fit for your organisation, consider:

- the virtual assistant may contain biases (for example, biases embedded in training data can influence how the AI perceives and responds to different genders)
- the privacy concerns that arise because virtual assistants often require access to meeting audio or video (the companion is essentially recording and analysing conversations)
- virtual assistants themselves can be vulnerable to hacking (if compromised, attackers could potentially access sensitive meeting information or disrupt the meeting flow with misinformation)
- virtual assistants can be expensive to implement and maintain and integrating this AI into existing workflows and ensuring user adoption can also be challenging
- over-reliance on virtual assistants can lead to decreased participant engagement and inattention (people might zone out if they expect the companion to handle tasks like note-taking or summarising discussions)
- virtual assistants can't replicate the human element of meetings (they currently lack the ability to understand emotional cues, build rapport, or facilitate sensitive conversations in the same way a human can) and this could lead to a less productive or emotionally sterile meeting environment

Also refer to our 'Tips for using AI tools' below.

The Governance Institute of Australia has published an issues paper 'Artificial Intelligence and board minutes – Issues Paper', which considers the potential applications and challenges of using AI in board meetings and company secretarial practice.

Content generation

Content generation is often a critical business function for many not-for-profits and charities as it increases outreach and engagement.

There are a range of AI tools (particularly generative AI) which can significantly enhance this aspect of an organisation's operations through generating various types of content – including social media posts and updates, community newsletters, annual reports, blogs and articles. In addition to written content, AI can be used to create images, infographics and videos.

Tools like ChatGPT and JasperAI which generate text content for articles, social media posts and newsletters can be coupled with others like Canva, Dall-E and Synthesia to create a full audio-visual content generation suite.

There are, however, a range of potential legal and ethical issues surrounding the use of AI generated content. These are discussed in more detail below and include whether content should contain disclaimers that it is AI-generated, copyright and IP infringement, and the potential to include misinformation.

Data analysis

Charities and not-for-profits can also use AI to organise and analyse large amounts of data for various sources and use the outputs to shape strategic direction and activities.

Insights generated through this process could be used to guide program development, measure the impact of activities and even identify funding and fundraising opportunities based on community and donor behaviour.

For example, there are a range of operational AI tools which can collect and analyse participant feedback from a particular workshop or activity, and then generate recommendations on how to improve it. Organisations can then use those recommendations to refine their strategies for greater impact.

Similarly, tools like DonationDrive AI and DonorSearch AI use an organisation's donor lists and donation data to predict when a donor is likely to give again or predict when someone is likely to become a donor for the first time. Organisations can use those predictions to tailor their outreach and target their fundraising campaigns to increase the likelihood of donations.

The main concerns regarding data analysis involve the potential disclosure of confidential and personal information. Publicly available AI models are often trained using the data which users input, meaning there is a risk that the information you share with an AI model becomes publicly available. This can have serious consequences and may mean it is inappropriate to use AI in some cases.

Aiding decision-making

Supported by data, AI can provide real-time data analysis which can be used to identify trends and inform strategic decision making.

✳ Example – use of AI for decision-making (managing the risks)

A local not-for-profit organisation that supports disadvantaged youth wants to use a grant management platform that incorporates AI functionalities to help allocate grant funding based on various criteria.

To proactively manage the risks of using AI, the organisation:

- conducts a bias assessment of the AI algorithm before deployment, ensuring it doesn't unfairly discriminate against certain group, and

- only uses the AI system as a decision-support tool, and ensures human experts review and approve the final grant allocations

> ✳ **Example – how Justice Connect uses AI**
>
> Every year millions of people miss out on legal help. Legal services funding isn't growing, but the huge amount of unmet legal need in our community is.
>
> To maximise our reach, Justice Connect uses AI in our online tools. These tools help people correctly identify their legal issue and connect with the right information and services.
>
> Justice Connect worked with the University of Melbourne to develop and train an algorithm to correctly identify legal issues based on the language of help seekers.
>
> We did this by analysing loads of data – specifically, the language people use to describe their legal issue (nearly 15,000 language samples) – and worked with pro bono lawyers to match that 'natural language' data with the correct legal issue (lawyers made more than 67,000 annotations). Using this data and analysis, Melbourne Uni built an algorithm.
>
> The model has been incorporated into our intake form. There's been a significant increase in people identifying their legal issue correctly, as well as an increase in conversion to services. We are now considering how else this technology could be used and applied.
>
> At the core of our innovation work, we have looked to develop safe, ethical and inclusive AI.

While AI offers significant potential for not-for-profit organisations, risks associated with use of AI raise ethical and legal concerns.

We discuss these below.

Legal risks associated with use of AI

> ❗ **Caution**
>
> The legal and regulatory landscape is still developing in Australia (and globally), as are the generally accepted practices for AI use. Organisations which use AI should keep themselves informed of any developments relevant to their AI use.

While there are no laws currently directly targeting and regulating AI across all sectors in Australia, existing laws apply to specific contexts of AI use in Australia. These include consumer protection laws, defamation laws, unlawful discrimination laws, WHS laws, intellectual property laws, and privacy laws.

Misleading information and defamation

AI-generated content (text, images, videos) could be inaccurate or misleading, causing harm to individuals or organisations. If an organisation publishes this content, they could be liable for spreading misinformation. If the content damages someone's reputation, it could be considered defamation.

In certain circumstances, the Australian Consumer Law (**ACL**) may apply. The provisions in the ACL are designed to protect consumers against behaviours which may be considered harmful to them, such as misleading or deceptive conduct, false or misleading representations, and unconscionable conduct.

If an organisation uses AI in a tool (for example, an AI-powered financial advice tool) that provides inaccurate recommendations, this could be seen as a defective service under the ACL.

For more information about the Australian Consumer Law, see our webpage 'Communications and advertising'.

For more information on defamation laws, see our webpage 'Understanding defamation laws'.

Intellectual property infringement

Use of AI could infringe intellectual property laws and rights, such as:

- **copyright** – it is not clear whether an AI program can be considered the 'author' of words it generates (and have its rights infringed) in Australia, but AI systems that generate content or use material protected by copyright owned by someone else, without proper authorisation, may infringe the other person's copyright

- **patent** – similarly, while Australian law does not recognise AI inventors (patents may only be granted for inventions by humans), an AI program which replicates or makes use of patented technology without permission from the owner may lead to patent infringement

- **trade mark** – AI-generated content or systems that use trade marks without permission may be liable for trade mark infringement

For more information on intellectual property infringement, see our webpage 'Understanding intellectual property laws'.

Data privacy and security

One of the most significant risks associated with the use of AI generally relates to data privacy and security.

Charities and not-for-profits often collect and store a range of personal information about their stakeholders, donors, beneficiaries and personnel. Depending on the area in which they operate, these organisations might also collect and handle sensitive or health-related information.

AI systems often require access to large amounts of data, and in many instances, this includes data of the organisation using the AI. In addition, many AI programs use the data which users input to continuously improve or 'train' themselves. This raises concerns about data privacy, and organisations need to ensure that their AI systems and use complies with privacy legislation including the Australian Privacy Principles.

Implementing AI tools which do not collect a user's data is one way of minimising the risk of privacy breach or, if that is not possible, limiting the types of information shared with AI tools. Notifying a person about how your organisation will use their personal information is a requirement under privacy laws. You should review your privacy policy and privacy statements or collection notices to make sure they include how you will or might share someone's personal information with AI.

For more information, see our webpage 'Privacy laws'.

Cyber security

AI systems can be targets for cyber attacks, leading to potential data breaches and security risks. In addition to liability for not implementing appropriate cyber security measures, some organisations may breach legislation like the *Security of Critical Infrastructure Act 2018* (Cth).

For more information, see our cyber security fact sheet.

Bias and unlawful discrimination

AI tools involve the use of algorithms which are 'trained' using large amounts of data. For this reason, there is a risk that AI will perpetuate existing biases or discrimination present in the data it is trained on. This can lead to unfair treatment and discrimination of certain groups or individuals.

For example, an AI tool used for assessing job applications might favour certain demographics if the training data reflects historic biases. Similarly, generative AI can produce content which is impacted by historical biases – some image generation AI will only show older men when asked to generate an image of 'company directors sitting around a board room table'.

Organisations should bear this in mind when using AI and may need to take active steps to combat bias such as removing personal and demographic information for job applications.

For more information on unlawful discrimination, see our resources on 'Managing people', including our webpage 'Discrimination laws'.

Lack of transparency

While AI can be very useful, not all AI systems are open about the way in which they make decisions. In addition, even though many organisations are now using AI, they do not always disclose how or when this occurs.

A lack of transparency as to AI's use within your operations can lead to a loss of trust among stakeholders, including donors and beneficiaries. Community organisations should ensure they are open about their use of AI, with clear explanations of how AI-driven decisions are made and what content is AI generated.

Job displacement

AI can be used to automate routine and repetitive jobs, as well as administrative tasks.

While this improves efficiency and can save costs, it can lead to job losses or significant changes in roles. Community organisations should balance the adoption of AI with strategies for upskilling staff and creating new roles that leverage human skills alongside AI capabilities, minimising negative impacts on the workforce.

Duty of care and work health and safety laws

Organisations have a duty of care under common law and work health and safety (**WHS**) laws. Using AI introduces potential legal risks related to this duty.

An organisation could be found to be negligent for failing to take reasonable steps to mitigate risks associated with AI if, for example, AI-generated decisions (such as grant allocation or service provision) lead to harm due to bias, misinformation or a data breach.

Under WHS laws, organisations have an overarching primary duty to ensure the health and safety (including psychological health and safety) of workers. Inadequate safeguards around use of AI could lead to psychological risks for staff, such as workplace anxiety or low employee morale and trust, for example due to:

- unfair decisions made by AI systems due to algorithm bias, or
- the threat of job losses due to AI automation

For more information on an organisation's duty of care, see <u>our resources on managing risk and insurance</u>.

For more information on an organisation's WHS obligations, see <u>our resources on work health and safety laws</u>.

Fraudulent and unlawful activity

AI can be used to facilitate fraud and other illegal activities, such as creating deepfakes or automating cybercrime like hacking and phishing attacks.

For example, an AI system for grant applications could be manipulated to generate fake requests. An organisation could be liable for financial losses resulting from such misuse.

Organisations should ensure that they and their personnel do not use AI systems for illegal purposes and implement measures to prevent and detect such misuse.

Compliance and legal risks

The regulatory landscape for AI is evolving, and charities and not-for-profits must stay informed about legal requirements and compliance issues. Non-compliance with laws and regulations related to use of AI can result in legal repercussions and damage to an organisation's reputation.

It is crucial to seek legal advice and implement compliance frameworks that address current and emerging regulatory requirements in Australia. These are addressed in more detail below.

Example – use of ChatGPT by an organisation's employee highlights risks

Sarah is employed by a not-for-profit organisation that provides education services. Excited about its potential to streamline content creation, Sarah uses ChatGPT to draft reports for the organisation's board and craft social media posts. The organisation doesn't have a policy on use of AI and Sarah doesn't tell her manager or the board that she has used ChatGPT in her work.

In the absence of transparency and established safeguards to ensure factual accuracy and avoid plagiarism, potential legal risks that Sarah's use of ChatGPT poses include:

- the inclusion of incorrect or misleading information in board reports or social media content because the data from ChatGPT may contain biases or factual inaccuracies
- infringement of intellectual property laws because the content generated may infringe copyright, and
- data privacy concerns if Sarah allowed ChatGPT to access the organisation's data

Sarah's use of ChatGPT could expose the organisation to legal liability. And the publication of misinformation or copyrighted material could damage the organisation's reputation with donors, partners, and the public.

Ethical concerns associated with use of AI

As AI plays a larger role in organisational decision-making, the issue of responsible and ethical use of AI is being considered across sectors and by governments.

The use of AI raises several ethical concerns, including issues related to consent, autonomy, and the potential for misuse.

For example, AI-driven surveillance tools might infringe on individuals' privacy rights, and automated decision-making systems might lack the nuanced understanding required for sensitive situations. Community organisations must navigate these ethical challenges carefully, establishing clear ethical guidelines and ensuring that the ways they are using AI align with their values and purpose.

The Australian Government has adopted a set of national AI Ethics Principles to promote safe, secure and reliable AI. The principles are voluntary and align with the international standards set out by the Organisation for Economic Development.

While the principles are not legally binding, the framework serves as a guide for responsible AI development and implementation in Australia.

The AI Ethics Principles comprise eight core principles for responsible use of AI:

1.   Human, social and environmental wellbeing
Throughout their lifecycle, AI systems should benefit individuals, society and the environment.

2.   Human-centred values
Throughout their lifecycle, AI systems should respect human rights, diversity, and the autonomy of individuals.

3.   Fairness
Throughout their lifecycle, 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
Throughout their lifecycle, AI systems should respect and uphold privacy rights and data protection, and ensure the security of data.

5.   Reliability and safety
Throughout their lifecycle, 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
Those 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.

The framework:

- is aimed at businesses and government agencies developing or using AI, and
- emphasizes human oversight, avoiding bias in AI systems, and ensuring data privacy

For more information, see the Australian Government's webpage 'Australia's Artificial Intelligence Ethics Framework.

Legal development in Australia

In recognition of the widespread use and increase in AI, the Australian Government released a discussion paper '*Safe and Responsible AI in Australia*' in June 2023. Over 447 published responses were received from a range of organisations (including Government, private sector, community organisations and individuals) and the Government published an interim response in January 2024.

The Government's published response gives an indication of the likely approach to AI regulation, but does not outline what that regulation might be. However, international efforts in Europe and Canada may provide some insight as to what Australia's approach might look like:

- The **European Union** approved the *Artificial Intelligence Act* (**AI Act**) in March 2024. The AI Act groups AI based on risk levels and imposes specific obligations for each category of either unacceptable, high, limited or minimal risk. The EU is considered a high watermark for AI regulation, which is in line with the approach it has taken in relation to privacy and the protection of personal information.

- **Canada**'s *Artificial Intelligence and Data Act* was first introduced in 2022 and is currently under consideration. If it goes ahead, an assessment and mitigation of risks for high-impact AI systems will be required and safety and ethical guidelines will be imposed. As with the EU, the Canadian approach would differentiate between levels of AI risks, requiring responses that reflect the specific nature and scope of potential impacts.

Based on the international experience, there is no 'one-size-fits-all' approach to AI, and it is likely that the Australian regulatory framework will recognise that different AI tools do different things, and the requirements need to be informed by what the AI is going to be used for.

For more information, see the Australian Government's discussion paper: Supporting responsible AI.

Note

Not-for-profit organisations using AI need to be aware of the legal risks associated with use of AI. Implementing responsible AI practices, including data governance, robust security measures, and human oversight, can help mitigate these risks and ensure AI is used ethically and legally.

How does AI impact organisational governance?

Directors, committee members and office-holders hold the ultimate responsibility for overseeing risk management throughout an organisation. This responsibility (which includes risk arising from AI) stems from a mix of legal sources – such as the duty of care and negligence, privacy laws and legal duties under common law and legislation.

Legal duties

Legal duties apply to people involved in governing a not-for-profit organisation, which includes people who sit on a committee or board, or are office holders (people with control and influence over the governance of an organisation, even if they don't hold an official position).

The legal duties that apply depend on the legal structure of an organisation (duties can apply to an unincorporated group) and whether the organisation is registered as a charity with the Australian Charities and Not-for-profits Commission (**ACNC**).

We have simplified the main duties by listing them as four main duties below. For more information about the legal duties that apply to your organisation, see our Duties guide.

> **In practical terms, a director, committee member or office-holder has the following four main legal duties:**
>
> 1. **Good faith** – the duty to act in good faith in the best interests of the organisation and for a proper purpose
>
> 2. **Reasonable care** – the duty to act with reasonable care, skill and diligence (including the duty to prevent insolvent trading)
>
> 3. **Conflicts of interest** – the duty to disclose and manage conflicts of interest
>
> 4. **Use of information or position** – the duty not to use the position (or information gained as a result of that position) improperly. This duty flows from the duty relating to conflicts.

An office-holder should not use AI for personal gain or for purposes that could undermine the public's trust in the organisation.

In complying with their duties, particularly the duties to act in good faith in the best interests of the organisation and with reasonable care, skill and diligence, office-holders should:

- understand how the organisation is using AI and manage the potential risks associated with those uses, and

- ensure AI is used ethically and aligns with the organisation's purpose

Note

At a minimum, directors and committee members of community organisations should be getting informed about AI, its uses, its benefits and shortcomings and the risks associated with its use.

Good governance and good decision making comes with awareness and understanding. By being informed about these matters, directors and committee members will be in a better position to be able to discharge their duties.

Directors and committee members should ask questions of management and others within their organisations about the use of AI, including:
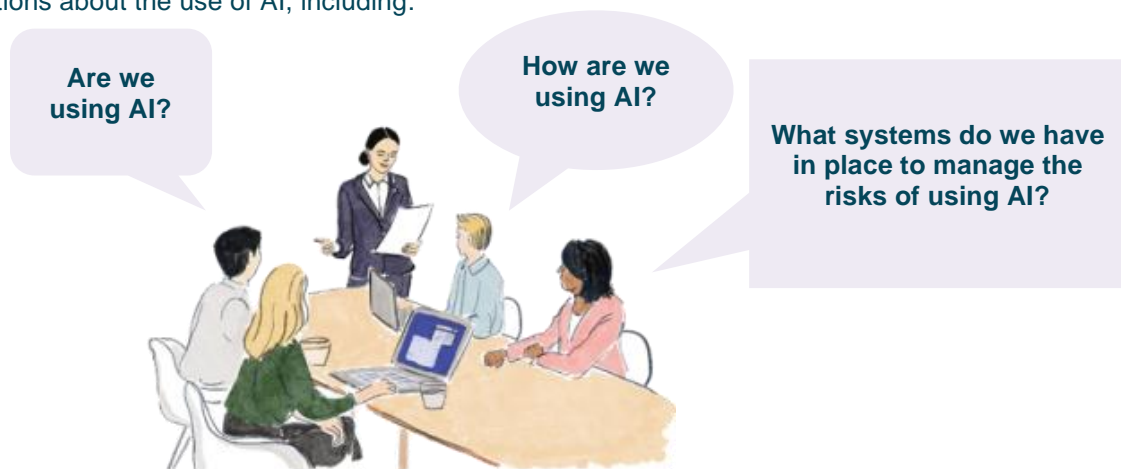
Armed with an understanding of AI and its use within their organisations, directors and committee members will be able to provide oversight and controls.

Implementing responsible AI practices will help to mitigate the risks associated with AI use identified above (for example, bias, discrimination, or data breaches) and protect the organisation from legal and reputational harm.

The Australian Institute of Company Directors has published a suite of resources on the use of AI, comprising:

- Director's introduction to AI
- Director's guide to AI governance
- a snapshot of the 'Eight elements of safe and responsible AI governance'
- a governance checklist

Also see the Institute of Community Directors Australia's help sheet: Artificial intelligence and governance frameworks.

Implementing responsible AI practices

To implement responsible AI practices, an organisation should:

- adopt an **AI risk assessment system** (or include AI as part of their usual risk management practices), which could include an AI oversight committee, to:
  - establish roles and responsibilities within the organisation around AI initiatives
  - keep up to date with legal and regulatory requirements and changing laws
  - conduct regular risk assessments for AI projects considering potential biases, transparency, data security, privacy concerns and misuse, and
  - develop appropriate mitigation strategies to address identified risks

- develop a comprehensive **AI policy** that establishes clear guidelines for responsible AI development and use in the organisation,

- where possible, prioritise AI systems with understandable decision-making processes to **promote transparency and trust**

- maintain **human oversight** and control over AI-driven decisions, particularly in high-risk scenarios

- integrate AI management with the organisation's existing policies and procedures to streamline governance and mitigate risks

- **educate staff** about AI capabilities, limitations, and potential biases to promote responsible use:
  - train staff on the organisation's AI policy and proper procedures for using AI tools
  - provide clear guidelines and policies for employees using AI tools for content creation or problem-solving, and
  - emphasize the importance of using AI ethically and responsibly, aligning with the organisation's values and purpose

Note – privacy laws

An organisation should already have a privacy policy that implements a framework for data collection, storage, usage, and security, complying with privacy laws.

For more information, see our webpage 'Privacy laws'.

Developing an AI policy

An AI policy should cover what is acceptable use of AI, (including data governance principles), set out risk mitigation strategies and human oversight procedures, and align with your organisation's existing policies including policies on data privacy, information security, and discrimination.

Information to include in an AI policy:	
Purpose and scope of the policy	• state the purpose of the policy (for example, to ensure the responsible use of AI within the organisation) • state the scope of the policy (for example, who does it apply to, what AI tools and uses are covered)
Principles to guide responsible use of AI	• **ethical use** – ensure AI technologies are used ethically • **transparency** – inform stakeholders about AI use and decisions • **accountability** – establish clear accountability for AI systems • **compliance** – adhere to relevant laws, regulations, and industry standards • **purpose and values** – AI use is consistent with the organisation's purposes and values and helps to achieve the organisation's purpose
Governance and oversight	• appoint a committee or officer to oversee AI activities (this person may also be responsible for exploring and implementing AI projects) • consider involving subject-matter experts and other stakeholders to provide input into AI activities and use • ensure the board or committee provides oversight of AI initiatives with regular formal and informal reporting
Risk management and data management	• conduct regular risk assessments for AI technologies depending on their use (identifying potential biases, security vulnerabilities, and privacy concerns) • implement processes for reducing the likelihood of risks (for example, through safeguards like bias detection tools, data security measures, and human review procedures) • implement measures to ensure data protection and compliance with the *Privacy Act 1988* (Cth) and Australian Privacy Principles, and any other applicable state or territory-based privacy laws • update the organisation's privacy policy
Ethical considerations	• ensure AI systems are free from bias and do not discriminate (and include this as a guiding principle in the governing body's and the AI committee's consideration of proposed AI use) • consider the impact of AI technologies on stakeholders like beneficiaries, employees and donors
Training and awareness	• provide staff and other key personnel with training on AI, its benefits, drawbacks and risks, and responsible use • provide training on the AI policy and its content and requirements
Monitoring and reporting	• implement continuous monitoring AI systems for correct and ethical functioning • provide regular reports on AI initiatives to the governing body and other stakeholders
Compliance and enforcement	• ensure all employees, volunteers, contractors, and partners comply with the policy • review the policy annually for relevance and effectiveness • outline consequences for non-compliance
Reviews and updates	• include a clearly defined process for reviewing and updating the policy at regular intervals, to ensure it remains current

Example – use of ChatGPT by an organisation's employee in accordance with an AI policy

Ben (the employee of a not-for-profit organisation) uses ChatGPT to draft reports for the organisation's board and craft social media posts.

In this scenario, the organisation has a policy on use of AI and Ben informs his supervisor about using ChatGPT and seeks guidance on appropriate content creation methods.

Ben and his manager review the organisation's AI policy to check that use of ChatGPT aligns with their guidelines. The AI policy includes guidelines on fact-checking and reviewing all AI-generated content before publication.

Ben follows the AI policy to ensure factual accuracy and to avoid plagiarism and other potential legal risks.

Tips for using AI tools

- Before using an AI tool, clearly define the problem you're trying to solve or the task you want to automate. This helps you choose the right tool with the functionalities that best suit your needs.

- Understand the AI tool's capabilities and limitations. Most AI tools have limitations in terms of accuracy, bias, and the type of data they can handle.

- If possible, choose AI tools that offer some level of explainability. This allows you to understand how the AI arrived at a particular decision or output.

- Be cautious about what data you feed into AI tools. Don't share information that belongs to your organisation, or information that is personal, sensitive, or confidential.

- Always follow your organisation's privacy and AI policies when using an AI tool.

- Ensure the words or data you use to prompt the AI tool are accurate, complete, and unbiased to avoid skewed or misleading results. The specificity of your prompts significantly impacts the quality of the outputs.

- Treat AI outputs as a starting point, not a final or correct piece. AI systems can be wrong, and their outputs might be biased or inaccurate. **Maintain human oversight and critical thinking when using AI-generated content or recommendations**.

- Always double-check information you receive from AI tools against reliable sources before using it for content creation or to make decisions.

- **Remember** – AI tools can inherit and amplify biases present in the data they're trained on. Be mindful of potential biases in your data and the AI tool itself. Consider using tools with built-in bias detection features.

- The field of AI is constantly evolving. Stay informed about advancements, best practices, and potential risks associated with AI tools.